BERNHARD BIERSCHENK

SJÄLVKONFRONTATION VIA INTERN TELEVISION I LÄRARUTBILDNINGEN

CWK GLEERUP / LUND

SJÄLVKONFRONTATION VIA INTERN TELEVISION I LÄRARUTBILDNINGEN

FÖRSÖK ATT MODIFIERA LÄRARKANDIDATERS PERCEPTION OCH VÄRDERING MED HJÄLP AV VIDEOBANDFÖRMEDLAD SJÄLVKONFRONTATION OCH TRADITIONELL HANDLEDNING

AKADEMISK AVHANDLING
som med vederbörligt tillstånd för vinnande av filosofie doktorsexamen
vid Lunds Universitets samhällsvetenskapliga fakultet
kommer att till offentlig granskning framställas i sal 122
å Lärrahögskolan i Malmö fredagen den 22 september 1972 kl. 10.00

AV

BERNHARD BIERSCHENK
FIL. LIC., VRML

Staffanstorps Tryckeri AB, 1972

BERNHARD BIERSCHENK

SJÄLVKONFRONTATION VIA INTERN TELEVISION I LÄRARUTBILDNINGEN

Staffanstorps Tryckeri AB, 1972

STUDIA PSYCHOLOGICA ET PAEDAGOGICA

SERIES ALTERA XVIII

Bernhard Bierschenk

Självkonfrontation via intern television i lärarutbildningen

FÖRSÖK ATT MODIFIERA LÄRARKANDIDATERS PERCEPTION OCH VÄRDERING MED HJÄLP AV VIDEO-BANDFÖRMEDLAD SJÄLVKONFRONTATION OCH TRADITIONELL HANDLEDNING

LUND / CWK GLEERUP

Till Inger

Förord

Föreliggande undersökning utgör en experimentell studie av lärarkandidaters reaktioner på sitt eget beteende i undervisningssituationer. Kommunikation och kontroll av information dels via dyadiska konfrontationsprocesser, dels via externt förmedlade konfrontationsprocesser med hjälp av intern television och videobandningsteknik har stått i centrum för undersökningen.

Genomförandet av denna studie har blivit möjligt tack vare ett samarbete mellan många olika människor — ett samarbete som har kännetecknats av ett ovärderligt engagemang hos de enskilda personerna.

Experimentet *Självkonfrontation via intern television i lärarutbildningen* har ingått som delundersökning i forskningsprojektet "Undersökningar av intern television, främst som utbildningsteknologiskt delsystem i en integrerad lärarutbildning", vars projektledare varit lektor Bertil Gran. Experimentet har utförts under vetenskaplig ledning av professor Åke Bjerstedt vid pedagogisk-psykologiska institutionen vid lärarhögskolan i Malmö.

Det är mig en mycket angenäm plikt att vid detta tillfälle tacka professor Bjerstedt inte bara för att han alltid har gett mig sitt stöd och visat mig nya vägar och infallsvinklar, utan även för att han skapat förutsättningarna för att jag, efter mina studier i Göttingen, grundligt fick lära känna och på beteendevetenskapliga grunder studera pedagogisk-psykologiska processer.

Även om författaren till denna skrift ensam är ansvarig för utformningen, genomförandet, utvärderingen och rapporteringen av undersökningen, har han under forskningsprocessens olika faser påverkats och stimulerats av den "ekologiska omgivningen". Den allmänt stimulerande forskningsatmosfären vid pedagogisk-psykologiska institutionen vid lärarhögskolan i Malmö, speciellt vid de integrerande forskningsseminarierna under ledning av professor Bjerstedt, och den generösa och förtroendefulla hållning som lektor Bertil Gran alltid har visat mot författaren, har varit väsentliga faktorer som har påverkat undersökningen.

Experimentets tekniska utformning och tidsbegränsning medförde att varje tekniskt missöde skulle ha spolierat projektet. Den ansvarige teknikern, Sten Solding, och institutionsteknikern Lennart Wramborg, som f ö har ritat figurerna, har genom stor noggrannhet och skicklighet kunnat hålla ITV-anläggningens inspelnings- och uppspelningsapparatur i funktion. Jag tackar både

Sten Solding och Lennart Wramborg för deras initiativrika och engagerade insatser, som i hög grad bidragit till ett friktionsfritt genomförande av experimentet.

I detta sammanhang vill jag också tacka lärarhögskolans rektor, fil dr Erik Stenquist, och utbildningsledarna lektorerna Lennart Hörberg och Arne Strelert för alla de ansträngningar och schematekniska tvärkastningar som har åstadkommits för att möjliggöra experimentets genomförande.

Dessutom vill jag framföra ett varmt tack för ett mycket gott samarbete till de lärarkandidater som har deltagit i experimentet som försökspersoner och som visat en mycket berömvärd punktlighet och stort tålamod.

Här vill jag särskilt tacka lektor Lennart Andersson. Hans medverkan har varit av mycket stor betydelse för experimentets genomförande.

Vid detta tillfälle vill jag också tacka dels lektorerna Ingrid Holmberg, Kurt Gestrelius, Per Sundgren och Lennart Wiechel för att de välvilligt har ställt sig till förfogande som panel för bedömning av experimentets protokollmaterial, dels folkskollärarna Sven Fransson och Lars Hansson, vilka som kodare har tagit hand om en stor mängd uttalanden av lärarkandidaterna.

Under experimentets olika faser har jag tacksamt emottagit råd av rektor Evert Lundberg och docent Bernt Larsson. Speciellt den senare vill jag tacka för alla våra dyadiska konfrontationer, under vilka det vuxit fram en mångfald av idéer kring undersökningsuppläggningar och lämpliga bearbetningsmetoder. Dessa dyadiska konfrontationsprocesser har haft välgörande effekter för den aktuella undersökningen.

Mitt tack för hjälp vid testrättningen och sorteringen av experimentets datamängder gäller herr Bengt Olof Ljungberg, fil kand Adela Zyto, men framför allt fil kand Marjanna Berg, som har haft huvudansvaret för denna arbetsfas.

Jag tackar fru Ingegärd Johansson för stansning av hela testmaterialet, fil kand Leif Robertsson vid Lunds Datacentral för genomförandet av de enskilda analysprogrammen och för utskrivning av manuskript fruarna Alice Pertzig, Ann-Kristin Larsson och Gunnel Bruhn.

Docent Anne Marie Wieselgren har utfört den slutliga språkgranskningen och Gillian Nilsson, B.A., har ombesörjt översättningen av sammanfattningen till engelska, för vilket jag tackar.

Slutligen vill jag uttrycka min stora tacksamhet för den omsorg och ihärdighet, som adjunkt Inger Ericson har visat vid åtskilliga språkgranskningar och korrekturläsningar under manuskriptens framväxt.

B.B.

Innehåll

DEL 2

Analyser av lärarkandidaters självbedömning 121

DEL 3

Analyser av pedagogiska experters bedömningar 159

DEL 1

BAKGRUND · EXPERIMENT · EXPERIMENTUPPLEVELSE

Inledning

För att studera lärarens och elevernas beteende i en undervisningssituation har många olika metoder för observation, registrering och evaluering utvecklats och använts. Dessa metoder baserar sig antingen på (1) direktobservationer (observatören är en del av undervisningssituationen) eller (2) indirekta observationer (någon teknisk anordning för registrering av audio- och/eller videosignaler har använts) eller (3) en kombination av dessa förfaringssätt. De observations- och registreringstekniker som hittills har kommit till användning inom skolforskningen avser nästan uteslutande den verbala interaktionen mellan lärare och elever. Detta faktum skulle bl a kunna förklaras med en vanlig uppfattning bland pedagoger, att interaktionen mellan människor huvudsakligen bygger på verbal kommunikation. Flanders påstår till exempel att den verbala kommunikationen utgör ett adekvat "sample" av lärarens totala påverkningsmönster (Flanders, 1965, s 252).

Att även den icke verbala interpersonella kommunikationen utgör en mycket viktig kanal för att förmedla information har emellertid uppmärksammats av bl a Birdwhistell (1953). Han menar att människor kommunicerar lika mycket med varandra via icke verbala tecken som via verbala. I en intervju (New York Times Magazine, May 31, 1970, s 42) säger Birdwhistell:

"Years ago I started with the question: How do body motions flash out words? Now I ask instead: When is it appropriate to use words? They're very appropriate to teach or to talk on the telephone but you and I are communicating on several levels now and on only one or two of them have words any relevance whatsoever. These days I put it another way: Man is a multi-sensorial being. Occasionally, he verbalizes."

Uppmärksamheten i den pedagogiska och psykologiska forskningen syns under senare tid alltmera ha blivit centrerad kring interrelationen mellan verbala och icke-verbala tecken (Biddle & Adams, 1967; Bellack, 1968; Brusling, 1971; Mortensen, 1971).

Den centrala problemställningen för den här presenterade studien innebär en granskning av möjligheten att använda intern television (ITV) och videobandning (VR) för registrering av audio-video-signaler och återgivning av omvärlden i rörelse i syfte att studera människan som "a multi-sensorial being". Studiet av i vad mån lärarkandidaten (lk) modifierar sin perception

och värdering till följd av videobandförmedlad självkonfrontation (en kombination av verbal och icke-verbal information) är förknippat med två fenomen, som berör individens personlighet på ett alldeles specifikt sätt, nämligen (1) att bli fotograferad och (2) att bli konfronterad med sitt eget ofta icke medvetna sätt att agera i olika situationer.

I vissa samhällen är det tabu att ta kort. Men även i vårt samhälle blir människor ofta pinsamt berörda, när man riktar en kamera mot dem. Den allmänna attityden mot fotografering är av tämligen emotionell natur.

Att bli konfronterad med sig själv är en perceptuell upplevelse av alldeles speciellt intresse för människan och upptar en central plats i hennes myter och fantasi. Myten om Narcissus dokumenterar detta väl. Robert Benchleys (1948, ss 117—120) beskrivning kan kanske tjäna som en inledning till problemställningen: *självkonfrontation via intern television i lärarutbildningen.*

"Merely as an observer of natural phenomena, I am fascinated by my own personal appearance. This does not mean that I am *pleased* with it, mind you, or that I can even tolerate it. I simply have a morbid interest in it.

Each day I look like someone, or some*thing,* different. I never know what it is going to be until I steal a look in the glass. (Oh, I don't suppose you really could call it stealing. It belongs to me, after all.)

One day I look like Wimpy, the hamburger fancier in the Popeye the Sailor saga. Another day it may be Wallace Beery. And a third day, if I have let my mustache get out of hand, it is Bairnsfather's Old Bill. And not until I peek do I know what the show is going to be.

Some mornings, if I look in the mirror soon enough after getting out of bed, there is no resemblance to any character at all, either in or out of fiction, and I turn quickly to look behind me, convinced that a stranger has spent the night with me and is peering over my shoulder in a sinister fashion, merely to frighten me. On such occasions, the shock of finding that I am actually possessor of the face in the mirror is sufficient to send me scurrying back to bed, completely unnerved.

All this is, of course, very depressing, and I often give off a low moan at the sight of the new day's metamorphosis, but I can't seem to resist the temptation to learn the worst. I even go out of my way to look at myself in store-window mirrors, just to see how long it will take me to recognize myself. If I happen to have on a new hat, or am walking with a limp, I sometimes pass right by my reflection without even nodding. Then I begin to think: 'You must have given off *some* visual impression into that mirror. You're not a disembodied spirit yet – I hope.'

And I go back and look again, sure enough, the strange-looking man I thought was walking just ahead of me in the reflection turns out to have been my own image all the time. It makes a fellow stop and think, I can tell you.

This almost masochistic craving to offend my own aesthetic sense by looking at myself and wincing also comes out when snapshots or class photographs are being passed around. The minute someone brings the envelope containing the week's grist of vacation prints from the drugstore developing plant, I can hardly wait to get my hands on them. I try to dissemble my eagerness to examine those in which I myself figure, but there is a greedy look in my eye which must give me away.

The snapshots in which I do not appear are so much dross in my eyes, but I pretend that I am equally interested in them all.

'This is very good of Joe', I say, with a hollow ring to my voice, sneaking a look at the next print to see if I am in it.

Ah! Here, at last, is one in which I show up nicely. By 'nicely' I mean 'clearly'. Try as I will to pass it by casually, my eyes rivet themselves on that corner of the group in which I am standing. And then, when the others have left the room, I surreptitiously go through the envelope again, just to gaze my fill on the slightly macabre sight of Myself as others see me.

In some picture I look even worse than I had imagined. On what I call my 'good days', I string along pretty close to form. But day in and day out, in mirror or in photograph, there is always that slight shock of surprise which, although unpleasant, lends a tang to the adventure of peeking. I never can quite make it seem possible that that is really Poor Little Me, the Little Me I know so well and yet who frightens me so when face to face.

My only hope is that, in this constant metamorphosis which seems to be going on, a winning number may come up sometime, if only for a day. Just what the final outcome will be, it is hard to predict. I may settle down to a constant, plodding replica of Man-Mountain Dean in my old age, or change my style completely and end up as a series of Bulgarian peasant types. I may just grow old along with Wimpy.

But whatever is in store for me, I shall watch the daily modulations with an impersonal fascination not unmixed with awe at Mother Nature's gift for caricature, and will take the bitter with the sweet and keep a stiff upper lip.

As a matter of fact, my upper lip is pretty fascinating by itself, in a bizarre sort of way."

1. Undersöknings- och rapporteringsprofil

1.1 Allmän problemställning

Lärares undervisningsskicklighet och personlighet betraktas som mycket väsentliga faktorer som påverkar undervisningsresultat (jfr kap 2). I den traditionella lärarutbildningen saknas dock uppenbarligen led mellan teoretiska beteendevetenskapliga kunskaper och övningar av tekniska färdigheter i en undervisningssituation. Tillgång till återkopplingstekniker skulle kunna underlätta integrationen av undervisningsupplevelser och teoretiska kunskaper. De senaste årens tekniska utveckling av ITV-system och videobandspelare tillåter nya forskningsmetodiska förfaringssätt för ett systematiskt studium av interaktionsprocesser. De hittills utförda undersökningarna i vilka denna teknik har tillämpats skulle kunna delas upp i

1. studier angående informationsselektion och beteendemodifikationer till följd av stipulerade beteendemodeller
2. studier angående informationsselektion och beteendemodifikationer till följd av individens predispositioner

En översikt över ett urval av den pedagogiska och psykologiska forskningslitteraturen rörande den audiovisuella självkonfrontationstekniken ges i kapitel 3. En målsättning för denna undersökning är att studera, i vilken utsträckning individens "själv" kan utnyttjas som kontroll och regleringsmekanism i syfte att utbilda lärare som kontinuerligt använder sig av återkopplingstekniker för att pröva, testa och modifiera det egna beteendet. "Själv" användes, med engelskans self som mönster, för att beteckna de erfarenheter, som individen gjort av personella och icke personella objekt och som antages ligga till grund för strukturen i individens personlighet (jfr Svenska Akademiens Ordbok, bokstaven S, sp 3 017).

1.2 Individpopulation

Målpopulationen för undersökningen är alla kvinnliga lkk med slutbetyg från gymnasiet som intogs vid landets lärarhögskolor ht 1967 och ht 1968 för ut-

bildning på mellanstadielärarlinjen. Uttagningen av det i experimentet ingående stickprovet samt inferensproblem vid generaliseringen från ett stickprov till individpopulationen diskuteras närmare i kapitel 4.

1.3 Design

I den pedagogisk-psykologiska forskningen har det sedan länge saknats metoder för att få den komplexa undervisningsprocessen under experimentell kontroll. Som forskningsmetodik har mikrolektionstekniken, en nedbantningsteknik, visat sig vara mycket användbar. Hela mikrolektionsstrukturen kan manipuleras för att man med hjälp av experimentella designer skall kunna besvara olika frågeställningar. I denna undersökning tillämpas en mikrolektionsteknik i samband med en faktoriell design för att besvara följande fråga: Vad betyder olika feedbackformer dels för individens modifiering av perceptions- och värderingsstrukturen (referensram), dels för individens beteendemodifikationer? Undersökningen omfattar två större avsnitt:

1. Ett strikt kontrollerat *experiment* där två feedbackformer (traditionell handledning och externt förmedlad självkonfrontation via ITV/VR) manipuleras.
2. En *långtidsuppföljning* sex veckor respektive fyra terminer efter slutförandet av experimentet.

Undersökningens experimentella utformning diskuteras utförligt i kapitlen 5 och 6.

1.4 Speciell problemställning

Undersökningens huvudproblemställningar är följande:

1. Vilka effekter har traditionell handledning (dyadisk konfrontation) respektive externt förmedlad självkonfrontation via ITV/VR på lkk självbedömning (perception och värdering)? Analys och diskussion av resultaten redovisas i del 2.
2. På vilket sätt percipierar och värderar pedagogiska experter lkk undervisningsprestationer? Expertbedömningar utgör ett objektivitetskriterium och indikerar i vilken utsträckning beteendemodifikationer kan konstateras. Analys och diskussion av resultaten redovisas i del 3.
3. I vilken utsträckning avviker lkk självbedömning från expertpanelens bedömningar? Analys och diskussion av resultaten redovisas i del 4.

En sammanfattande beskrivning av experimentets design, resultat och implikationer avseende dessa tre problemställningar samt några rekommendationer redovisas i del 5.

4. Experimentets huvudinstrument granskas bl a med hjälp av kanonisk korrelationsanalys. Analys och diskussion av resultaten redovisas i del 4.
5. En granskning av samband mellan experimentets resultat och sådana bakgrundsvariabler som är närmare specificerade genom testbatteriet (se kap 8) planeras att redovisas vid ett senare tillfälle.
6. En granskning av långtidsuppföljningen redovisas vid ett senare tillfälle.

En mera detaljerad redovisning av de frågor som experimentet skulle kunna ge svar på finns i kapitel 2.

Tillvägagångssättet vid strukturering, organisering och lagring av experimentets olika datamängder för en datorbaserad behandling redovisas i kapitel 9. En analysplan för utvärderingen av de observerade data i enlighet med experimentets problemställningar presenteras i kapitel 10. Metodiska problem i samband med strukturering av olika datanivåer med hjälp av faktoranalyser har redovisats i kompendium nr 12 (Bierschenk, 1971 b).

1.5 Teoretisk förankring

Experimentet är av explorativ karaktär, och dess huvudsyfte är att bidra till upptäckandet av nya relationer och nya insikter i den komplexa undervisningsprocessen för att man skall kunna formulera forskningsproblem mera precist och för att man skall kunna utveckla mera strikt formulerade hypoteser. Men hur explorativ en undersökning än må vara, måste det finnas någon teoretisk förankring. I annat fall kan man inte definiera undersökningens mätområde eller testa hypoteser. Denna undersökning bygger på ett fenomenologiskt synsätt som presenterades i samband med utvecklingen av ett kategorisystem (Bierschenk, 1971 a), på vilket konstruktionen av experimentets mätinstrument har baserats. Det fenomenologiska synsätt som presenterades av Combs & Snygg (1959) har bl a varit grundläggande för denna undersökning. Författarna konstaterar (s 17):

> "People do not behave according to the facts as others see them. They behave according to the facts as they see them."

Den tekniska innebörden, som här ges begreppet "själv," implicerar, att individens personlighet är uppbyggd av inlärda *själv-andra-* och *själv-icke personella objekt*-relationer. Relaterad till experimentet kan utgångspositionen formuleras på följande sätt: (1) Lk har möjligheten att kunna se sig själv som ett "*objekt*", som är uppbyggt av en mängd olika erfarenheter och (2) lk beteenden påverkas endast när informationen har någon mening för individen ifråga, dvs när information korresponderar med individens "själv". Om själv betraktas som en organisation av erfarenheter, så borde detta innebära, att förändringar av individens beteenden förutsätter förändringar i individens "själv". (För en mera ingående diskussion hänvisas till Bierschenk, 1971 a, ss 25—29.)

Syftet med experimentet är alltså att studera individens *perceptions- och värderingsstruktur* samt dessas reorganisation eller omstrukturering till följd av den experimentella påverkan. Individens sätt att percipiera och värdera olika objekt, dvs hans subjektiva tolkning av en given situation och handling, har under senare tid blivit föremål för definitioner inom ramen för den moderna systemforskningen, där man alltmera tagit intryck av den grundläggande principen, att alla informationsprocesser är urvalsprocesser. Analysen av individens tillvägagångssätt och effektivitet vid val, kategorisering och lagring av tillgänglig information, står numera i uppmärksamhetens centrum (Buckley, 1968, ss 120—121). Detta synsätt har medfört att intresset inte längre enbart koncentreras på olika former av information och på inlärningens struktur utan i allt större utsträckning på kommunikationens effektivitet.

Ett grundläggande drag i systemforskningen är den kontinuerliga utvecklingen av allmänna modeller som bygger på feedback och kontroll, dvs man utgår ifrån att kybernetiska principer ligger till grund för utvecklingen av komplexa adaptiva system, som människan utnyttjar för sin egen anpassning till omvärlden. De första tillämpningarna av kybernetiska modeller i beteendevetenskapligt sammanhang har redan visat sig vara framgångsrika (Heinich, 1968, s 108). Den kybernetiska modellen kan bl a tjäna som förklaring av självreglering och självkontroll via feedbackmekanism och har således en direkt relevans för experimentet. I ett tidigt bidrag till diskussionen om feedbackproblemet i samband med social diagnos och analys av handlingsmönster presenterar Lewin redan 1947 feedbackmekanismen som ett instrument för självreglering (Lewin, 1968, ss 441—444).

De kybernetiska principer som framfördes av Miller m fl (1960) har varit vägledande i experimentet. När en undervisningsprocess analyseras med utgångspunkt från en kybernetisk modell, innebär detta att de olika komponenterna i modellen inte reagerar på varandra enbart på ett mekaniskt pingpong-liknande sätt, utan att de var och en på ett tämligen aktivt sätt bidrar till utvecklingen av den pågående undervisningsprocessen. Varje enskild individ interpreterar situationen mot bakgrund av sin egen referensram. Med ledning av denna interpretation väljer eller bestämmer sig individen för en viss handling. Detta förutsätter användning av *planer* eller *strategier* oberoende av hur fragmentariska och oslipade eller logiska och genomarbetade dessa är.

Perception betyder enligt en mycket allmänt hållen definition en interaktion eller transaktion mellan individ och omgivning. I denna undersökning avses med perception en sammankoppling av den interna strukturen som används som test med de konkreta beteenden eller handlingsmönster som kom till användning i den aktuella situationen.

Värdering definieras som ett val eller en viktning av olika alternativa beteenden. Värdering refererar till bestämda beteendeaspekter och indikerar vilken betydelse olika undervisningsbeteenden har för individen ifråga.

1.6 Variabelpopulation och konstruktion av mätinstrument

Experimentets mätinstrument konstruerades i syfte att mäta karakteristiska drag i individernas perception och värdering av en klassrumssituation.

Beslutet om vilka items som skall ingå i ett mätinstrument måste ske på godtyckliga grunder. Men för att ge impulser till urvalet av items genomfördes ett förförsök, där lkk fick till uppgift att ge spontana och simultana kommentarer under externt förmedlade självkonfrontationsprocesser via ITV/VR. Dessa kommentarer blev föremål för en innehållsanalytisk bearbetning (se Bierschenk, 1972 f). Kategorierna i denna utgjorde sedan basmaterialet för konstruktionen av de enskilda items. Mätinstrumenten och mättillfällena diskuteras i kapitel 8, medan de validitets- och reliabilitetsproblem som är förbundna med mätinstrumenten diskuteras i del 2 och 3.

1.7 Pedagogisk-psykologiska problemställningar och hypotesprövningar

Ett huvudinstrument i undersökningen är skattnings- och värderingsformulär F III, som i enlighet med kategorisystemet innehåller sex olika a priori definierade *subjekt-objekt-relationer,* som har definierats i samband med kategorisystemets konstruktion. En kortfattad definition av de enskilda relationerna ges nedan:

1. *Ego-ego-relation*
   Perception och värdering av attribut (egenskaper, relationer) som gäller den egna personen *utan* att någon annan person är inblandad.

2. *Ego-elev-relation*
   Perception och värdering av attribut avseende annan person (elever).

3. *Ego-IPO-relation (dvs relationer mellan den egna personen och icke-personella objekt)*
   Perception och värdering av omgivningen: klassrummet, tekniska anordningar samt lektionsförberedelse, val av undervisningsmetod, mera formella metodiska aspekter, t ex tidsuppskattning, tempo i undervisningen etc.

4. *Elev-ego-relation*
   Perception och värdering av elevernas explicita eller implicita operationer mot lk.

5. *Elev-elev-relation*
   Perception och värdering av elevernas operationer sinsemellan respektive elevens agerande gentemot sig själv, där lk står utanför interaktionsprocessen.

6. *Elev-IPO-relation (dvs relationer mellan elev och icke-personella objekt)* Perception och värdering av elevernas operationer mot omgivningen: klassrummet, teknisk apparatur, ämnet etc.

I experimentet skall testas huruvida lkk perceptions- och värderingsstruktur förändras påvisbart i förhållande till de ovan angivna 6 subjekt-objekt-relationerna genom de olika påverkningskombinationerna. Frågeställningen kan alltså formuleras om till följande hypoteser:

H_0: traditionell handledning ger ingen effekt
H_0: självkonfrontation via ITV/VR ger ingen effekt
H_0: samspel mellan traditionell handledning och självkonfrontation via ITV/VR föreligger ej.

1.8 Inre och yttre validitet

Syftet med ett experiment är att kunna generalisera resultatet till en individgrupp och till sådana situationer, som ej ingår i experimentet. Olika generaliseringsproblem och generaliseringsmöjligheter i samband med experimentella undersökningar diskuteras i kapitel 11. I samma kapitel redovisas också några empiriska resultat som belyser vissa aspekter av experimentets ekologiska validitet, dvs i vad mån experimentet har upplevts som en exceptionell förteelse jämfört med den vanliga utbildningssituationen.

2. Bakgrund och syfte

All utbildnings målsättning är att påverka individer i syfte att åstadkomma beteendemodifikationer. En av förutsättningarna för sådana beteendemodifikationer är enligt författarens mening, att utbildningen för individen resulterar i en förändring av det perceptuella fältet. Med ett sådant synsätt blir undervisningen starkt relaterad till personligheten, eftersom denna involveras i den studerandes sätt att uppfatta sig själv och de erfarenheter som han kontinuerligt gör (Combs & Snygg, 1959, s 367). Betraktas individens personlighet som ett system, uppbyggt av olika erfarenheter, så borde man kunna påverka individens beteenden genom att tillföra systemet nya erfarenheter.

Om det kan utvecklas ett ekonomiskt och praktiskt användbart system som visar sig kunna effektivt påverka individer, borde detta medföra viktiga konsekvenser för den nuvarande lärarutbildningen. Ett sådant system skulle kunna vara av betydelse för lärarhögskolor och andra utbildningsanstalter.

Lärarens undervisningsskicklighet och personlighet betraktas a priori som mycket väsentliga faktorer som påverkar undervisningsresultatet (Malmquist, 1969, s 90). Undersökningar i vilka lärareffektiviteten granskats som en funktion av elevernas beteende- och prestationsförändringar och där man studerat i vad mån lärarens självperception och värdering i undervisningssituationer kan användas som ett reglerings- och kontrollinstrument för att åstadkomma "önskade" lärarbeteenden är än så länge förhållandevis få.

En ökning av lärarens praktiska undervisningserfarenhet behöver enligt Swineford (1963, ss 214—224) ej nödvändigtvis garantera att undervisningen blir bättre. Han visade att det kan vara vilseledande att tro, att ett ökat antal år av undervisningserfarenhet på fältet medför en större undervisningsskicklighet. Hälften av de lärare ($n = 500$), som undersöktes, visade en undervisningsstil som de övade in under utbildningstiden.

Ett resultat i annan riktning visar dock Malmquists (1969, ss 93—96) undersökning, där han försökte bestämma sambandet mellan elevers lässkicklighet och lärarens (1) levnadsålder, (2) antal tjänsteår, (3) slutbetyg från seminariet och (4) slutbetyg i undervisningsskicklighet.

Resultatet av sambandsprövningen ($\alpha = .01$) visade att lärare med 12

tjänsteår i genomsnitt åstadkom betydligt bättre resultat (beroende variabel: läsindexet) än lärare med mindre än 12 tjänsteår. Även ett χ^2-test ($\alpha = .05$) avseende skillnaden mellan lärare med 6 och fler respektive 6 och färre tjänsteår blev signifikant. Malmquists undersökning visar bl a ett svagt samband mellan lässkicklighet och levnadsålder och antal tjänsteår samt nollkorrelation mellan lässkicklighet och slutbetyg i undervisningsskicklighet.

Innebörden i både Swinefords och Malmquists undersökningar är dock ej lätt att tolka. De resultat som presenteras av Heil, Powell & Feifers (1960, ss 68—81) belyser mera direkt på vilket sätt lärarens beteenden påverkar elevernas prestationer. Författarna relaterade nämligen elevernas prestationer till interaktionen mellan lärar- och elevpersonligheten. Tre lärarpersonlighetstyper och fyra elevpersonlighetstyper identifierades med hjälp av ett antal kognitiva test och personlighetstest. De olika lärare-elev-kombinationerna jämfördes på basis av elevprestationer och lärarens kunskaper samt klassrumsobservationer. Den välintegrerade (självkontrollerande) läraren (typ A) visade sig vara mest effektiv med avseende på alla elevtyper, medan den dåligt integrerade (ängslige) läraren (typ B) visade sig vara ineffektiv på alla elevtyper, dock med undantag av "strebertypen". Den turbulente läraren (typ C) definierades av författarna som den lärare som liknar den defensive, intellektuelle naturvetaretypen, dvs intellektualitet uppfattas här som försvarsmekanism för den vuxna personens ego. Denna typ av lärare visade sig vara framgångsrik i att undervisa sådana elever som karakteriserades som konformitets- eller strebertyper. Dessa lärare uppnådde även mycket bra resultat i matematik och inom de naturvetenskapliga ämnena. Men de var ineffektiva vid undervisningen av ständigt opponerande eller obeslutsamma elevtyper. Sådana elever kräver en förmåga till interpersonella kontakter som dessa lärare visade sig sakna.

En bra undervisningsförberedelse och en bra undervisning kräver en kontinuerlig uppmärksamhet på de problem som är involverade i lärarens självkognition (perception och värdering) och strävan att förbättra sin undervisning.

En viktig uppgift i lärarutbildning, vidareutbildning och fortbildning är en kontinuerlig förbättring av lärarens undervisningsbeteende, inte minst med hänsyn till de krav som ställs på grund av den ständigt pågående reformverksamheten i skola och samhälle.

Ett centralt problem i lärarutbildningen är därvid att integrera utbildningens teoretiska och praktiska moment. Bjerstedt (1966, ss 179—180) konstaterar:

"Tyvärr bildar teoriundervisning och praktiska övningar ibland jämförelsevis slutna världar, inom vilka det talas delvis skilda tungomål. Dessa förhållanden gör det ej lätt för lärarkandidaten att spåra sammanhangen mellan den teoretiska bakgrunden och det praktiska handlandet."

Hur litet som egentligen har hänt i lärarutbildningen sedan John Deweys dagar kan antydas genom en jämförelse mellan detta uttalande och Deweys påstående att:

"The student adjusts his actual methods of teaching, not to the principles which he is acquiring, but to what he sees succeed and fail in an empirical way from moment to moment . . . In this way the controlling habits of the teacher finally get fixed, with comparatively little reference to principles in the psychology, logic and history of education" (Dewey, 1962, s 7).

Dessa förhållanden kan dels bero på att det trots alla ansträngningar under de senaste decennierna finns alltför få metoder för systematiska undersökningar av lärarbeteendet och undervisningsprocessen, dels på en rad markanta svårigheter i den traditionella lärarutbildningen (för diskussion se Bjerstedt, 1966, ss 179—184).

I den pedagogiska litteraturen poängteras mycket kraftigt vikten av observationsmöjligheter i lärarutbildningen. Observationernas värde härledes framför allt ur behovet att åstadkomma en integration av de teoretiska och praktiska aspekterna i lärarutbildningen. Många lärarutbildare medger, att auskultationer och övningsundervisning i den nuvarande utformningen i mångt och mycket är otillfredsställande och inte leder till önskade beteendemodifikationer. Det vore alltså önskvärt att förbättra för närvarande tillämpad praxis.

Ett stort antal undersökningar har tidigare utförts för att beskriva och analysera lärarens reaktioner, när det önskade undervisningsbeteendet var känt och accepterat av läraren. Denna undersökning har däremot till syfte att studera förändringar när det *ej finns i förväg externt och explicit fastställda normer.*

Jersild (1955, ss 13—14) poängterar, att det under de senaste åren genom ett ökat antal undersökningar allt mera blivit påtagligt, att självförståelse kräver helt annorlunda metoder än att studera undervisningsplaner och -uppgifter. Lärarutbildningen är i mycket stor utsträckning baserad på just sådana moment som av den enskilde lk kräver, att denne upptäcker de moment i undervisningsprocessen som i hög grad är beroende av lk personlighet.

Lk behöver alltså någon form av feedback, för att det över huvud taget skall kunna inträffa en förändring av lk perception och värdering, som i sin tur kan leda till en modifikation av hans egen referensram vilket bl a är en förutsättning för beteendemodifikationer. För denna undersöknings ändamål definieras "modifikation av perceptions- och värderingsstrukturen" som förändringsgraden i lk perception och värdering av de genom mätinstrumenten specificerade beteendeaspekterna.

På det pedagogiska fältet har intern television (ITV) och videobandning (videorecord, VR) samt mikrolektioner (ml) alltmera kommit i uppmärksamhetens centrum. I debatten uttrycks förhoppningen att en mängd problem

skall kunna lösas med hjälp av ITV. Många skolor och högskolor har utrustats med kompletta ITV-anläggningar, dock utan att man samtidigt tagit hänsyn till de utbildningsteknologiska processerna och de inbegripna uppgifterna, dvs utan att man försökt åstadkomma en analys och syntes av de komponenter som är relaterade till ett sådant system.

De senaste årens tekniska utveckling av ITV-system och videobandspelare har öppnat nya vägar för att utveckla metodiska förfaringssätt, baserade på reglerings- och kontrollprinciper (t ex fördröjd audiovisuell feedback), som vidgar våra möjligheter att systematiskt undersöka lärarens sätt att observera interaktionsprocesser, i huvudsak relaterade till hans eget agerande i undervisningssituationer. Även användningen av mikrolektioner som en nedbantningsteknik (Allen & Ryan, 1969, s 2) bör anföras i sammanhanget, då denna teknik ger både ökade möjligheter att åstadkomma en integration av teoretiska och praktiska aspekter i lärarutbildningen och större möjligheter att få den komplexa undervisningsprocessen under kontroll både i utbildnings- och forskningssammanhang. En hörnsten i lärarutbildningen världen runt syns vara handledda auskultationer och övningsundervisning. Praxis är att lk i någon övnings- eller praktikskola observerar en erfaren lärare för att sedan bl a försöka imitera dennes undervisningsfärdigheter, -strategier och -tekniker. Även om det förutsättes att lk har lärt sig någonting (imitation kan ju vara en möjlig inlärningsform), så är regleringsmöjligheterna för lärarkandidaten i stor utsträckning begränsade till handledarens responser i egenskap av "förstärkare". En svårighet i den traditionella handledningssituationen, då man vill hjälpa lk att upptäcka sig själv och sina beteendebenägenheter, ligger i handledarens begränsade möjligheter att på ett exakt sätt kunna rekonstruera de händelser som inträffat i övningsundervisningen. Denna svårighet kan emellertid elimineras med hjälp av de nya media.

Lärarutbildaren har all anledning att uppmärksamma den interna televisionens och videobandningens speciella användningsmöjligheter som utbildningsinstrument i handledningssituationer. Att låta lk studera sig själv, dvs *auskultera på egna lektioner* är inte längre ett pedagogiskt önsketänkande utan en i realiteten lätt tillgänglig teknik. Därvid ställer man sig självfallet följande frågor: (1) Hur upplevs denna självkonfrontation via ITV i lärarutbildningssituationen, och vilka effekter har den? (2) Hur fungerar den jämfört med den traditionella handledningen? och (3) Skulle individens "själv" kunna fungera som ett kontroll- och regleringsinstrument för att kontinuerligt kunna åstadkomma en förbättrad undervisning?

Ett förförsök. För att få en första uppfattning om individens reaktion vid självkonfrontation via ITV/VR genomfördes under juni månad 1968 ett explorativt förförsök med endast sex lkk. Försöket var främst avsett att ge en uppfattning om på vilket sätt lkk upplevde och kommenterade självkonfrontationsprocessen och på vilket sätt en större undersökning borde utformas

(lektionslängd, tidsintervaller, antal uppspelningar, apparatanordning etc). För att inte påverka utvecklingen av försökssituationen och för att få så mycket information som möjligt, styrdes försöket i ytterst liten utsträckning. Bland datainsamlingsmetoderna användes "simultankommentarer" för att ge en föreställning om i vilken riktning lkk reaktioner skulle kunna gå. Instruktionen till lkk i samband med simultankommentarerna var att kommentera självkonfrontationen samtidigt med uppspelningen av mikrolektioner. Att utan en sådan studie utveckla en design för en större undersökning och konstruera ett mätinstrument skulle kunna ha medfört dels risk för onödiga felkällor och dels risk för att påverka lkk med ett urval av items som föga motsvaras av den experimentella situationen.

För att på ett mera systematiskt och kontrollerat sätt kunna studera vilka effekter externt förmedlad självkonfrontation via ITV/VR har, jämfört med den traditionella handledningsmodellen, på lk perception och värdering dels av sig själv, dels av undervisningsprocessen, utformades det i denna rapport presenterade experimentet. Målsättningen för experimentet är bl a

1. att studera effekten på utvalda beteendeaspekter genom externa auskultationer på *egna lektioner* förmedlad via ITV och VR
2. att studera eventuella förändringar i lk perceptionsfält och värdering av den egna undervisningen vid *olika* undervisnings- respektive uppspelningstillfällen med *varierande* tidsintervaller
3. att studera lk självbedömning och -skattning av den egna undervisningen i en mikrolektion i *relation* till en pedagogisk expertpanels skattningar.

Experimentet ingår, som nämnts i förordet, som en delundersökning i ett forskningsprojekt som år 1967 påbörjades av professor Åke Bjerstedt och lektor Bertil Gran vid Pedagogisk-psykologiska institutionen vid Lärarhögskolan i Malmö. Projektet planerades bl a för att granska en rad specialfrågor i vilka ITV-arrangemang ingår som delsystem i en integrerad lärarutbildning (t ex utnyttjande av självkonfrontationsteknik och konstruktion av videobandade beteendetest av simulatortyp etc. För utförligare beskrivning hänvisas till Bjerstedt, 1967 a och b, 1970).

Det här presenterade experimentet är, som redan sagts, av explorativ karaktär. Syftet med en sådan explorativ studie är att vinna förtrogenhet med ett fenomen eller att förvärva insikter för att mera precist kunna formulera forskningsproblem och kunna utveckla hypoteser. Huvudsyftet är alltså att upptäcka nya relationer och att få nya insikter. Detta förutsätter att forskningsplanen utformas tillräckligt flexibelt, så att ett fenomen kan studeras ur många olika aspekter. Denna studie är ett första försök att med hjälp av en experimentell uppläggning studera lk informationsselektion samt förändringar i perceptions- och värderingsstrukturen, när lk blir konfronterad med sin egen prestation i en undervisningssituation. Av de frågor som vi hoppas få svar på

genom experimentet kan följande nämnas. (Inom parentes har placerats sådana problem som icke upptas här, men för vilkas behandling data finns lagrade.)

1. *Lärarkandidater: självbedömning*
   1.1 I vilka avseenden ger handledning och självkonfrontation via ITV/VR skilda effekter?
   1.2 Finns det samspelseffekter mellan traditionell handledning och självkonfrontation förmedlad via ITV/VR-tekniken?
   1.3 I vilken utsträckning sker en förändring i lkk perception och värdering vid en konfrontation med nytt videobandmaterial?
   1.4 I vilken utsträckning är lkk perception och värdering (a) ego-centrerad, (b) elev-centrerad eller (c) stoff-centrerad?
2. *Bedömningspanel: expertbedömning*
   2.1 Vilken grad av samstämmighet uppnås inom en pedagogisk expertpanel vid bedömning av mikrolektioner?
   2.2 Skiljer sig enligt expertpanelen lkk undervisningsprestationer?
   2.3 Kan expertpanelens data utnyttjas som jämförelsekriterium, t ex som skattning av lk relativa *"objektivitet"* vid bedömning av den egna personen och undervisningen?
3. *Lärarkandidater — bedömningspanel*
   3.1 I vilken utsträckning förändrar den experimentella påverkan "objektiviteten" i lkk självbedömning?
   3.2 Finns det i pedagogiska experters och lkk perceptioner och värderingar dimensioner, som inbördes är relaterade till varandra och kan dessa läggas till grund för konstruktionen av tolkbara index?

(4.) *Samband mellan experimentets resultat och sådana bakgrundsvariabler, som är specificerade genom testbatteriet som presenteras längre fram*
   (4.1) I vilken mån finns det samband mellan typiska perceptionstendenser och testdata?
   (4.2) I vilken mån kan avvikelserna från sambanden enligt föregående punkt relateras till andra variabler, t ex kognitiva test, personlighets- och attitydtest?

(5.) *Långtidsuppföljning*
   (5.1) I vilka avseenden skiljer sig lkk bedömning av de egna videobandade mikrolektionerna sex veckor efter slutförandet av experimentet, dels sinsemellan, dels i förhållande till tidigare bedömning?
   (5.2) I vilka avseenden skiljer sig lkk bedömning av de egna under experimentet videobandade mikrolektionerna 24 månader efter slutförandet av experimentet, dels sinsemellan, dels i förhållande till de tidigare utförda bedömningarna?

3. Pedagogisk och psykologisk forskningslitteratur rörande den audiovisuella självkonfrontationstekniken

Självkonfrontation via ITV/VR har under de senaste åren blivit en populär teknik och lett till ett stort antal rapporter och artiklar av mycket varierande kvalitet. Dels har rapporterna publicerats i många olika tidskrifter, dels föreligger ett antal icke tryckta doktorsavhandlingar. De flesta rapportförfattarna skriver entusiastiskt om ITV/VR-teknikens många inneboende möjligheter. Alger & Hogan (1966, s 1) skriver t ex att det ej borde vara någon överdrift att påstå att upptäckten av ITV/VR-teknikens användningsmöjligheter inom psykiatrin till sin betydelse kan jämföras med upptäckten av mikroskopets användningsmöjligheter inom biologin.

Att Alger & Hogans påstående, representativt för många liknande, verkligen återspeglar en realistisk bedömning av ITV/VR-teknikens användbarhet som pedagogisk-psykologiskt forsknings- och utbildningsinstrument återstår dock att bevisa. Det finns ett rikt antal känslomässiga påståenden i de forskningsrapporter, där ITV/VR utnyttjats för att studera självkonfrontationsprocessen, men mig veterligt inga undersökningar i vilka man, på objektiva grunder, har kunnat demonstrera *"entydigt" mätbara relationer mellan självkonfrontation, förmedlad via audio-visuella hjälpmedel, och personlighetsförändring.* Undersökningar har ofta utförts med bristfälliga designer, och dessutom kommer det stora flertalet av rapporterna från den klinisk-psykologiska verksamheten. Många studier som granskats i samband med vårt experiment kännetecknas tyvärr av en hel rad experimentella fel såsom avsaknad av en basgrupp (gruppen utan påverkan), brister i urvalet av försökspersoner (fpp) eller ej lämpliga och/eller otillräckliga mätinstrument. De mätinstrument, som utnyttjats, har nämligen vanligen utformats för andra syften än för respektive undersökning. Innan självkonfrontationstekniken kan testas mera ingående, bör alltså konfrontationsmetodiken utformas på basis av ett antal strikt kontrollerade experiment. Utgående från ett antal grundforskningsresultat skulle man sedan friare kunna laborera med den mångdimensionella självkonfrontationsvariabeln.

För att illustrera den hittills utförda forskningen rörande den audiovisuella självkonfrontationstekniken ges nedan några resultat som nästan uteslutande har rapporterats i doktorsavhandlingar. Resultaten är av motsägande natur och bör därför ses som en subjektiv information som enbart kan ge vissa ledtrådar för det fortsatta forskningsarbetet. Med hänsyn till undersökningsmetodiken kan forskningsrapporterna indelas i två stora grupper nämligen studier som tillämpar någon form av givna normer (3.1) och studier som undersöker fpp predispositioner och dessas påverkan på informationsselektion och informationsvärdering (3.2).

3.1 Studier angående informationsselektion och beteendemodifikationer till följd av stipulerade beteendemodeller

Till studier angående stipulerade beteendemodeller kan sådana räknas som har utförts vid *Stanford Center for Research and Development in Teaching* och vid *USA Air Force*. Så visar bl a McDonald & Allen (1967) och Eachus (1965) konsistenta beteendemodifikationer hos lärare och officerare. Dessa modifikationer har åstadkommits med hjälp av självkonfrontationstekniken.

Acheson (1964) undersökte feedback via enbart ITV, feedback via ITV i samband med olika handledningsmetoder samt feedback utan varje handledning. Beroende variabel är i vilken utsträckning fpp håller monologer respektive åstadkommer ett samtal med eleverna, dvs aktiverar elever till egna bidrag till undervisningen. I ”*direkt*” handledning gav handledaren konkreta förslag till beteendemodifikationer, vid ”*indirekt*” handledning fick fpp ej något direkt förslag till beteendemodifikationer. I stället försökte handledaren hos fpp framkalla föreställningar om på vilket sätt de trodde att undervisningsprestationen skulle kunna förbättras. Denna påverkningsform ledde ej till signifikanta resultat på 5 %-nivån. Acheson kommer till slutsatsen att VR som feedbackmetod kombinerad med handledning är ett användbart instrument i en handledningssituation för att underlätta fpp analys och förändring av det egna beteendet i en undervisningssituation. Resultaten blev signifikanta med $\alpha = .05$ och $\alpha = .01$. Med hjälp av VR förbättras också i begränsad utsträckning handledarens möjlighet att genom såväl direkt som indirekt handledning åstadkomma beteendemodifikationer hos fpp, vilket framgick av efterföljande videobanduppspelningar.

Fanslow (1967) undersökte ”35 mm Time-Lapse Photography” som ett feedbackinstrument för att göra fpp (lkk) medvetna om i vilken utsträckning eleverna är uppmärksamma respektive ouppmärksamma. Någon signifikant förändring ($\alpha = .05$) i fpp perception kunde dock ej påvisas. Fanslow rapporterade emellertid ett intressant resultat. Under experimentet visades histogram på vilka fpp perception jämfördes med flera av varandra oberoende bedömares (12 highschool-elevers) perception. Vid flera tillfällen riktades fpp uppmärksamhet på den stora diskrepansen mellan fpp bedömning och de för fpp okända bedömarnas skattningar. Fpp kommentarer var då att bedömaren ej kände till eller förstod individen ifråga lika bra som fp själv eller att bedömaren otvivelaktigt skulle ha bedömt annorlunda, om han själv hade befunnit sig i samma situation. Som framgår av Fanslows diskussion förklaras detta resultat med att de oberoende bedömarna av fpp ej upplevts som någon ”fara”, eftersom bedömarna ej hade något reellt inflytande på fp, och om detta skulle ha varit fallet, så var det en eventuell fara som lätt kunde åsidosättas av individen (Fanslow, 1967, ss 61—62).

Walz & Johnston (1963) undersökte vilken effekt självkonfrontation har på självperceptionen hos ”counseling candidates” (studerande som utbildas för psykiatrisk eller psykologisk rådgivning). Undersökningen visade, att fpp efter att ha sett sig själva lättare accepterade andra personers bedömningar av de egna yrkesfärdigheterna samt att de blev mindre positiva i självvärderingen. Undersökningen visade också att videobandning erbjuder ett unikt instrument för att under praktiktiden hjälpa de studerande att förstå sig själva bättre och vara mera medvetna om patienternas känslor. Författarna observerade en stor variation i fpp reaktioner, men återupplevelsen av ett terapeutiskt moment hjälpte många att förändra perceptionen utan direkt hjälp av handledaren.

Ward (1965) undersökte ”Interpersonal Process Recall” (IPR)-tekniken i samband med användningen av videobandspelare och ljudbandspelare. Denna teknik ger deltagaren i en dyadisk konfrontation, omedelbart efter den avslutade konfrontationen, ökade möjligheter att dels återge minnesbilder, dels återge upplevelser från den dyadiska konfrontationen. Parallellt, men isolerade från varandra, återges sedan den på ljud- och videoband upptagna konfrontationen. Vid återgivningen ställer en tredje person frågor rörande vissa väsentliga moment i den dyadiska konfrontationsprocessen. Frågorna är avsedda att underlätta återupplevelsen och tolkningen av beteenden. Genom simultana frågor kan man få parallella reaktioner från båda individerna som har varit involverade i den dyadiska konfrontationsprocessen. Undersökningen resulterade dock ej i signifikanta differenser ($\alpha = .05$), vilket förklaras av Ward med att mätinstrumentet var för grovt för att kunna ge utslag.

Unruh (1967) genomförde en metodstudie för att kartlägga olika registreringsmetoders effektivitet vid efterföljande skattning av undervisningssituationen genom en bedömningspanel. I undersökningen jämfördes maskinskrivna utskrifter, ljudbandinspelningar, videobandinspelningar samt en kombination av dessa. Bl a undersöktes också på vilket medium bedömarna baserade sina ”korrekta” bedömningar av lärarens effektivitet enligt ett ”givet effektivitetskriterium”. Sammanfattningsvis kan sägas att den mest valida källan för de givna stödstimuli, som ligger till grund för bedömarnas skattningar, syns vara audio-videobandning av lärarbeteenden. Alla metoder, som användes för ana-

lysen, visade att de audio-visuella inspelningarna var de bästa källorna för skattningen av lärarens effektivitet. En förklaring till hur detta kan vara fallet är bl a att den audio-visuella presentationen mest liknar den vanliga klassrumssituationen (Unruh, 1967, s 115).

Young (1967) undersökte femton olika metodiska aspekter som ingår i lkk utbildning i metodik i syfte att bestämma den relativa effektiviteten vid övning av speciella undervisningsfärdigheter. Lkk fungerade som modellärare och hade till uppgift att konkretisera dessa metodiska aspekter. Lektionsmodellerna videobandades och interfolierades med kommentarer. Grundmodellen som Young (1967, ss 116—117) diskuterar i resultatanalysen är en tvåfaktor-modell. Den första faktorn avser tillfällig respektive systematisk fokusering på de femton metodiska aspekterna. Den andra faktorn representerar "specifika" illustrationer respektive presentation av en "fullständig" lektionsmodell. Någon signifikant huvudeffekt har ej kunnat påvisas, men en tillfällig fokusering tycks vara den lämpligare träningsmodellen, eftersom en tillfällig fokusering i kombination med "specifika" illustrationer tycks vara överlägsen en systematisk fokusering i kombination med en "fullständig" beteendemodell.

Orme (1966) visade däremot att det är mera effektivt att (1) handleda genom en fysisk person än att (2) konfrontera lk med hjälp av videobandade undervisningsmodeller eller också att (3) konfrontera lk med de egna videobandade lektionerna.

Frågan huruvida en handledare i en dyadisk konfrontation med lk handleder mera effektivt, dvs förmedlar undervisningsstrategier och -handlingar mera framgångsrikt än det är möjligt dels via videobandade beteendemodeller, dels via en konfrontation med det egna undervisningsbeteendet förmedlad via ITV/VR är dock hittills ej besvarad.

Abbott (1965) undersökte möjligheten att använda självkonfrontationstekniken. Vid början av en behandlingsrelation mellan terapeut och patient spelade han in patientens beteende i en behandlingssituation. Syftet med denna studie var att undersöka effekterna på patienternas beteende när de konfronteras med det egna beteendet. Den slutsats som kan baseras på statistiskt signifikanta resultat gällande de effekter som självkonfrontationen hade på behandlingsprocessen är: Självkonfrontationen resulterade i början i en minskning av det ansvar som patienten visade inför behandlingsprocessen. Självkonfrontationen hindrade dock ej patienten från ett ansvarstagande vid slutet av experimentet. Kontrollgruppen visade mindre variationer i beteende och tog ansvar på samma sätt. Abbotts slutsats är att experimentgruppens responser på självkonfrontationserfarenheterna antydde, att konfrontationen med det egna beteendet resulterade i ett ökande intresse för den egna personen samt att den ökade individens motivation till förståelse av det egna jaget (Abbott, 1965, s 111).

De Lemos' (1963) syfte var att studera effekten av fördröjd återkoppling

(delayed feedback) på intervju-beteenden hos en grupp elever i klass 7. Antagandet var att information i form av experimentella betingelser skulle kunna hjälpa till att öka självförståelsen, definierad som högre självexploreringsnivå hos eleverna. I samband med elevernas självexplorering undersöktes bl a försvarsmekanismer. Undersökningen ledde ej till påvisbara effekter på avsett område. Däremot visade sig en oväntad tendens: Den mera erfarne handledaren framkallade högre självexploreringsvärden än den mindre erfarne.

Oliveros (1964) problemställning gällde bl a följande frågor: (1) Kan videobandning av lkk beteenden i en undervisningssituation ersätta direktobservationer i lärarutbildningen? Mikrolektioner användes vid datainsamlingen. Det undersöktes bl a om feedback från en handledare (supervisor), som observerade videobandinspelningar, åstadkommer flera förändringar i fpp beteenden än feedback, som kommer från en handledare, som observerade lektionen direkt i klassrummet. (2) Är det nödvändigt att få feedback från en handledare för att kunna förändra ett beteende? (3) Åstadkommer verbal feedback från en handledare i samband med videoband-feedback flera förändringar i fpp beteenden än *enbart* verbal feedback från en handledare? Och slutligen: (4) Ger en handledare, som kommer från ett universitet (university supervisor), större påverkan än en handledare, som kommer från en skola (school supervisor), mätt i form av större förändringar i fpp beteenden? Undersökningen visade, att fpp som fick någon form av feedback förändrade beteendet, definierat som "utveckling av mål" och "lärar-elev-kommunikation", mera än de fpp som ej fick feedback. Olivero drar slutsatsen att fpp behöver veta hur bra prestationen har varit, om en förändring i deras beteendemönster skall inträffa.

Fpp som fick tillfälle att se sig själva samt att få verbal feedback från en handledare, visade större modifikationer i beteende, som definierades som "förståelse av mål", jämfört med fpp som enbart fick verbal feedback. Oliveros slutsats är att såväl fpp som handledaren utnyttjar videobandningen för att etablera en "common frame of reference", som gör det lättare för handledaren att åstadkomma en konstruktiv kritik. I vissa fall blev fpp ej medvetna om att de hade betett sig på ett speciellt sätt, förrän handledaren hade påpekat detta speciella beteende på monitorn.

Den experimentella betingelsen att observera undervisningsbeteende på en monitor innan handledning påbörjades visade sig ej vara överlägsen direktobservationen, dvs handledarens närvaro i klassen. Skillnaderna visade sig dock vara systematiska och gick i den predicerade riktningen. Fpp som fick feedback från en handledare från ett universitet, blev påvisbart bättre på kriteriet "förståelse av mål" än de fpp som fick handledning av handledare från en skola. Det fanns även positiva förändringar (dock ej statistiskt påvisbara) i "utveckling av mål", "lärar-elev-kommunikation", "uppspelning som en del av värderingen" och "förstärkning som en del av värderingen".

Angående frågan, om videobandning skulle kunna vara ett substitut för direktobservationer i lärarutbildningen, kommer Olivero till resultatet att fpp uppenbarligen behöver feedback för att kunna åstadkomma beteendemodifikationer. Verbal feedback plus video-feedback hade större effekter än enbart verbal feedback. Visserligen kan man av dessa resultat ej utan vidare dra slutsatsen att videobandning kan vara ett substitut för direktobservationer i lärarutbildningen. Å andra sidan tillåter resultaten dock slutsatsen att vissa utvalda beteendemönster kan förändras med hjälp av olika feedback-betingelser. Användningen av videobandning kan mot denna bakgrund ses som ett effektivt instrument för att förändra beteenden i önskad riktning (Olivero, 1964, s 82).

Stroller (1967) fann att schizofrena fpp, efter en viss tid av självobservation av sin fysiska apparition, började iakttaga sitt eget icke-önskvärda beteende och försökte att ändra detta. Liknande resultat visar undersökningar som utnyttjade andra metoder för informationsgivning i fråga om det egna beteendet. Gage, Runkel och Chatterjee (1960) undersökte lärarens förändrade beteende i klassrummet som följd av negativ information från eleverna. Resultatet visar att beteenden förändrades i den riktning som den negativa informationen implicerade. Författarna baserar sina prediktioner på kongruens-teorin, som i enlighet med relaterade balans-teorier (kongruitet och dissonans) skulle kunna användas för att undersöka "feedback" i sociala sammanhang. Zajonc (1960, ss 280—296) förde en integrerande diskussion kring likheterna mellan olika balansteoretikers modeller. Det kan mot denna bakgrund antas att lärare, som ser en videobandning av sin egen lektion, vilken visar sig vara inkongruent med de värderingar som läraren har (t ex läraren ger uttryck för vikten av att eleverna deltar i diskussioner och kommer sedan att observera, att han för det mesta talar själv, dvs håller monolog), kommer att förändra sitt eget beteende i framtida undervisning.

3.2 Studier angående informationsselektion och beteendemodifikationer till följd av individens predispositioner

Den andra gruppen av studier utgöres av sådana, där specifika standardbeteenden (normer) varken är givna eller accepterade. Det måste alltså finnas andra variabler som bestämmer personens reaktioner på självkonfrontationsprocessen. I dessa studier antages, att reaktioner på information som t ex kommer från videobandningar, film eller ljudbandspelare, är beroende av iakttagarens självkognition och predisposition, vilka leder till att specifika informationskategorier accepteras. Bl a påpekade Wylie (1961, s 246) och senare Steiner och Rogers (1963, ss 128—136) att den huvudsakliga motivationen för att selegera ny information om det egna beteendet ej ligger i strävan att uppnå

kognitiv balans. I stället syns motivationen vara den bakomliggande variabeln i strävan att öka eller åtminstone bibehålla "self-esteem". Wolf (1943) och Nielsen (1962) rapporterar extrem emotionalitet och "rejection" samt undvikande av den information som personen fick genom att se sig själv via film. Detta är ett mycket viktigt och intressant resultat som vi har skäl att återkomma till.

Nielsens fpp fick diskutera sin egen livsfilosofi med en "tränad" intervjuare. Mot slutet av diskussionen provocerade intervjuaren medvetet fp för att se på vilket sätt fp kontrollerade sig själv medan fp var tvungen att försvara sin livsfilosofi. Nielsens studie visade bl a, att fpp i en efterföljande uppspelning av en dyadisk konfrontationssituation ej var i stånd till opartisk självvärdering, eftersom situationen fortfarande var känsloladdad. Efter 18 månader, när fpp återigen konfronterades med den inspelade situationen, kunde de flesta fpp lösgöra sig, vilket betyder, att självvärderingen blev mera balanserad. Vid denna slutliga konfrontation visade sig ett mycket intressant resultat. De fpp, som i början gav uttryck för en positiv självvärdering, reagerade nu negativt och de fpp som i början visade en negativ självvärdering värderade nu sig själva mera positivt. Nielsen konstaterar sammanfattningsvis att självkonfrontationsmetoden hade en välgörande effekt på fpp. Fpp var faktiskt i stånd att iaktta de egna kroppsliga reaktionerna, lyssna till den egna rösten och värdera innehållet i diskussionen. De fick se sina nervösa och andra emotionella uttryck. Slutligen kunde fpp observera på vilket sätt förhållandet till en annan person utformades, medan den egna livsfilosofin försvarades (Nielsen, 1962, ss 167—168).

Kerber (1967) försökte bestämma den omfattning och de påverkningseffekter som konfrontationen med lkk eget undervisningsbeteende har på deras sätt att utveckla ett "personal-professional concept" i en undervisningssituation. Som mätinstrument användes i denna undersökning "Cantril Self-Anchoring Scale". Undersökningen (Kerber, 1967, ss 106—108) visade bl a tre former av förändringar hos lkk:

1. en förändring från ett idealistiskt till ett mera realistiskt sätt att se undervisningen
2. en förstärkning av känslan av att vara nybörjare samt en ökad medvetenhet om de svårigheter som är förknippade med att uppnå "lärarídealet"
3. en mera kritisk och mera analytisk inställning än den som fpp visade innan de fick se sig själva.

Alla fpp trodde att en kontinuerlig självvärdering är nödvändig för att åstadkomma en förändring från nybörjare till framgångsrik lärare och att det alltid finns ett behov *att analysera den egna undervisningen.*

Pepperman (1967) undersökte den effekt som videobandning i "counselor trainee"-utbildningen har på kandidaternas självvärdering utan förmedlande påverkan av en handledare. Kandidaternas intervjubeteende videobandades. Resultaten (Pepperman, 1967, s 75) blev bl a att de patienter, som behandlades av den traditionellt utbildade kandidatgruppen, skattade dessa kandidater mera positivt än de patienter, som tillhörde den grupp av kandidater, som fick se sig själva. Kandidaternas självskattning vid slutet av experimentet visade ej någon signifikant skillnad mellan båda grupperna. Gruppen som fick se sig själv visade under experimentets gång en signifikant förbättring av objektiviteten i självskattningen, medan den traditionellt utbildade gruppen från början till slut ej visade någon förändring i självskattningen. En analys av speciella beteendemönster visade beteendemodifikationer. Gruppen som fick videobandning visade en signifikant förbättring i "positiva kommunikationsbeteenden". Ingen förändring kunde påvisas för någon av de båda grupperna beträffande "negativa kommunikationsbeteenden", nervösa manér samt positiva och negativa icke verbala beteenden.

Urvalet av handledare spelar en väsentlig roll för utbildningen av lkk, eftersom handledarens sätt att se handledningsprocessen också påverkar handledningen. Pepperman (1967, s 19) anför i sin litteraturöversikt att handledning är kognitiv och informationsgivande men med negativa övertoner. För att studera detta påstående, handleddes en "counselor trainee"-grupp "affektivt" och en grupp "kognitivt". De båda grupperna skilde sig påvisbart från varandra. Skillnaderna beskrivs dock tyvärr ej närmare.

Roulx (1967) undersökte effekten av ljudbandspelare, ljudbandspelare och strukturerad handledning samt strukturerad handledning utan ljudbandspelare på "counselor-in-training". För att objektivt kunna mäta dessa påverkningseffekter användes fysiologiska mått som t ex kardiogram och hudtemperaturmätning. Analysen visade ej någon signifikant differens i kandidaternas fysiologiska reaktioner under de olika påverkningsformerna. Resultatet av undersökningen kan sägas vara att inspelningen av handledningssektionen i och för sig ej producerade ångest hos kandidaterna, när de informerades om att ljudbandet ej skulle ligga till grund för handledningen. Användes ljudbandspelare i syfte att tjäna som underlag vid handledningen förändrades dock betydelsen av denna metod. Ljudbandinspelningar avsedda att användas vid senare handledning hade ångestutlösande effekter hos kandidaterna.

Walz & Johnstons undersökning (1963, ss 232—236) visade, att den "obehagliga" upplevelsen att se sig själv ej förhindrar en beteendemodifikation. Att iakttaga sig själv medförde tvärtom beteendeförändringar i den riktning som implicerades i de givna meddelandena. Undersökningen tyder på att självobservation via ITV/VR förändrar kandidatens perception utan en förmedlande påverkan av en handledare (jfr Pepperman, 1967, s 19). Av allt att döma är personens reaktion på att se sig själv påverkad av attityder mot honom själv och av i vilken utsträckning personen i fråga är nöjd med sin egen prestation samt av andra personspecifika predispositioner som t ex ångest. Å andra sidan kunde Kagan, Krathwohl och Miller (1963, ss 237—243), som använde videobandning som hjälpmedel för att stimulera återupplevelse ("recall") i terapeutisk behandling, visa att lkk perception gick mera i riktning mot handledarens perception än i riktning mot elevernas perception.

Salomon & McDonald (1969) undersökte informationsselektion och attitydförändringar hos lkk som för första gången såg sig själva i en undervisningssituation. Undersökningens resultat visade att självaikttagelse leder till attitydförändringar som är relaterade till fpp predisposition. I denna undersökning kunde dock ej påvisas det faktum, som visats i många andra undersökningar, att de flesta självobservationerna blir negativt värderade av fpp. Självvärderingen medförde ej någon förändring i negativ riktning. Salomon & MacDonald (1969, s 14) konstaterar att när det ej finns någon modell för "good teaching", ej någon handledning och när ej nya och allmänna normer är accepterade, är reaktionen på att se sig själv och sin egen prestation i en undervisningssituation i stort sett bestämd av fpp predisposition. Detta betyder att fpp belåtenhet med sin egen prestation före uppspelningen kommer att bestämma vad som observeras på en TV-monitor, på vilket sätt detta värderas och till vilka attitydförändringar det leder. Tyvärr utnyttjade Salomon & McDonald ej någon kontrollgrupp-design, t ex en grupp som fick handledning i samband med att se sig själv. Resultaten kan därför ej direkt jämföras med andra studier, där bl a presentationen av nya normer ("standards") för beteenden och anpassning till sådana har utnyttjats.

4. Experimentets genomförande och urval av försökspersoner

Experimentet (se fig 2, s 44) genomfördes vid Pedagogisk-psykologiska institutionen vid Lärarhögskolan i Malmö (LHM). Tekniska och personella begränsningar ledde till en uppdelning av experimentet i två faser. Experimentets båda faser genomfördes i lärarhögskolans ITV-studio och två angränsande laboratorier.

Under vårterminen 1969 utfördes den första fasen. Därvid genomfördes mikrolektioner uppföljda av traditionell handledning och mikrolektioner uppföljda av extern självkonfrontation via ITV/VR. Under vårterminen 1970 utfördes experimentets andra fas, nämligen en kombination av mikrolektioner uppföljda av självkonfrontation via ITV/VR *och* traditionell handledning samt enbart mikrolektioner.

För en undersökning om och i vilken utsträckning lärarutbildningen har påverkat lkk perception och värdering samt om och i vilken utsträckning de experimentella betingelserna medfört långtidseffekter planerades att konfrontera lkk under slutterminen med de under andra terminen videobandade mikrolektionerna (se fig 3, s 63). De lkk som vårterminen 1969 studerade andra terminen vid LHM avslutade sin lärarutbildning vårterminen 1971 (6:e terminen).

För de lkk som deltog i experimentets andra fas genomfördes denna del av experimentet vt 1972.

4.1 Urvalsförfarandet

Syftet med en experimentell undersökning är att kunna generalisera de vunna resultaten från det i experimentet ingående stickprovet av individer till en individpopulation, dvs att kunna dra slutsatser om undersökningsresultaten med hänsyn till en större grupp än den som ingick i det aktuella experimentet.

Denna generalisering kan utföras i två inferenssteg:

1. Från det stickprov av individer som ingår i experimentet dras inferenser till den för experimentet tillgängliga populationen.

2. Från den tillgängliga populationen generaliseras resultaten till målpopulationen.

Den tillgängliga populationen är definierad som den population, som experimentatorn har tillgång till, medan målpopulationen är definierad som den totala grupp av individer om vilken man genom ett experiment vill öka sina kunskaper (Bracht & Glass, 1968, s 440). Har uttagningen av lkk till experimentet skett genom ett slumpförfarande och är den tillgängliga populationen väl definierad, innebär första inferenssteget vanligen ej något större problem. Det är då enbart en fråga om vilken inferensstatistik som kommer till användning.

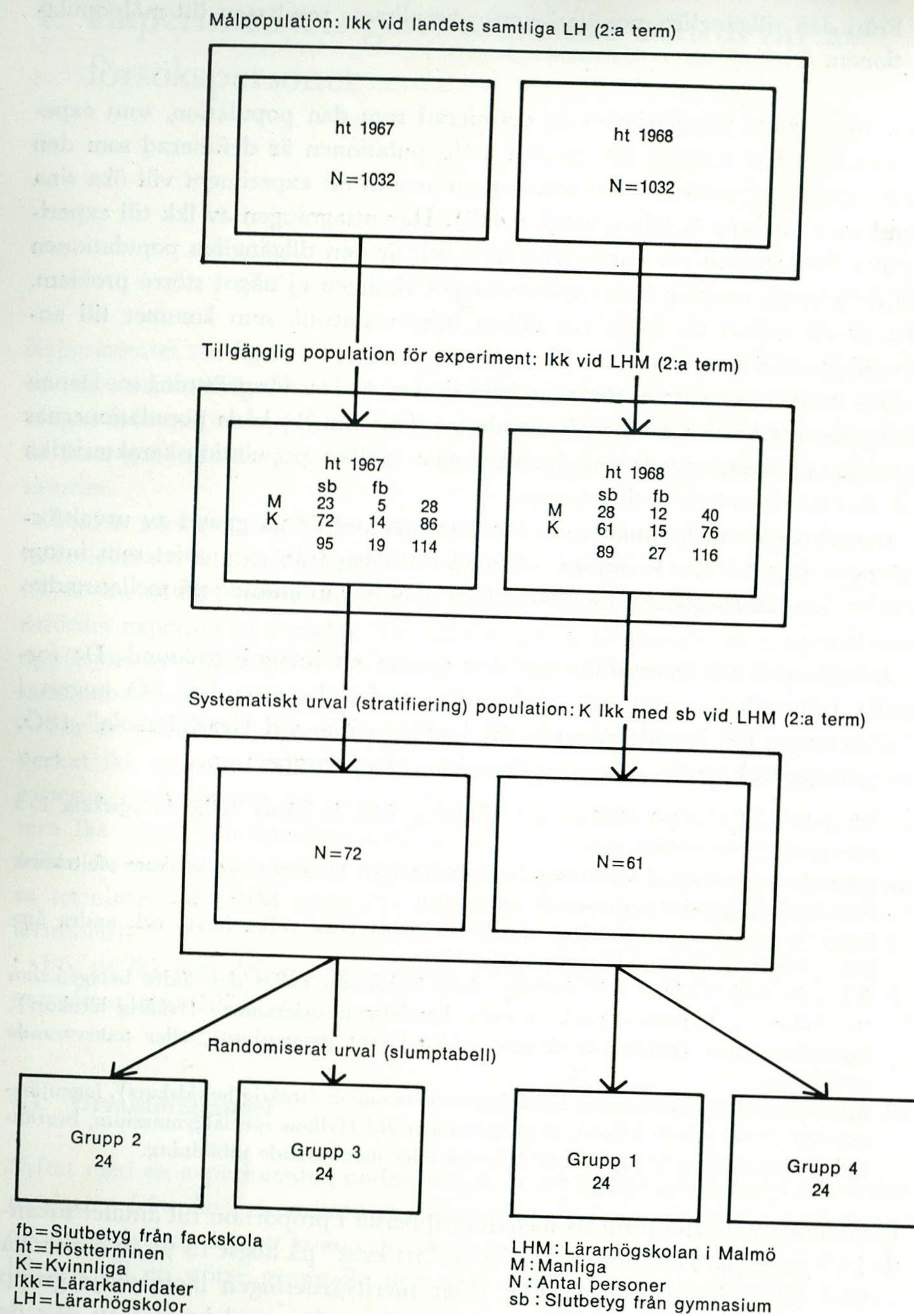
Det andra steget kan göras under mindre strikta förutsättningar. Denna generalisering kräver dock en ingående kunskap om (1) båda populationernas speciella egenskaper och (2) interaktionen mellan populationskarakteristika och de experimentella betingelserna.

Den speciella målpopulationen för experimentet är på grund av urvalsförfarandet (fig 1) alla kvinnliga lkk med slutbetyg från gymnasiet som intogs vid landets lärarhögskolor ht 1967 och ht 1968 för utbildning på mellanstadielärarlinjen.

Intagningen till lärarutbildning sker genom en intagningsnämnd. De formella behörighetskraven och nödvändiga förkunskaperna har SÖ angett i "Vägledning för inträdessökande till klasslärarlinje vid lärarhögskola" (SÖ, Stockholm, 15.4.1969). Alla sökande indelas i fem grupper:

a) Sökande med slutförd lärokurs och slutbetyg från fackskola enligt betygsskala 1–5 eller motsvarande utbildning.
b) Sökande med slutförd lärokurs och slutbetyg från gymnasium eller kurs på teknisk linje i gymnasiet eller motsvarande utbildning.
c) Sökande med tredje årskursen från folkhögskola eller sådan första och andra årskurs i folkhögskolan som bygger på grundskolan.
d) Sökande med lärokurs och slutbetyg från fackskolan enligt den äldre betygsskalan (berömlig, . . .), normalskolekompetens, handelsgymnasieexamen (tvåårig lärokurs), ingenjörsexamen (tvåårig heltidskurs vid tekniskt gymnasium), eller motsvarande utbildning.
e) Sökande med studentexamen, handelsgymnasieexamen (treårig heltidskurs), ingenjörsexamen (treårig heltidskurs), avgångsexamen vid Hvilans specialgymnasium, begränsad studentexamen vid försvarets läroverk eller motsvarande utbildning.

Utbildningsplatserna fördelas mellan grupperna i proportion till antalet sökande i varje grupp. Dessutom finns det en "fri kvot" på högst 15 % av det totala antalet utbildningsplatser på LH. Efter meritvärderingen inom varje grupp rangordnas de sökande och företrädet bestäms efter medelvärdet av dels de betyg som ingår i slutbetyg, dels de betyg över fyllnadsprövning eller särskild prövning i ämne som ej ingår i slutbetyg men som fordras för behörighet. De studerande som ingår i den fria kvoten kan få tilläggspoäng med högst .05



Figur 1. Uttagning av lärarkandidater till experimentet

poäng för annan utbildning, som bedöms vara av betydelse för lärarverksamhet. Med ledning av rangordningen inom respektive behörighetsgrupp beslutar en intagningsnämnd om preliminär antagning. Beslut om definitiv antagning fattas av rektor för den lärarhögskola dit den studerande har sökt.

Experimentet utgick från den på lärarhögskolan i Malmö tillgängliga populationen av approximativt 240 lkk, som intogs vid lärarhögskolan ht 1967 och ht 1968 för utbildning på mellanstadielärarlinjen. Den tillgängliga populationen består alltså av två terminskurser. För inferensdragningen hade det varit mera tillfredsställande om den tillgängliga populationen hade varit identisk med målpopulationen, dvs alla lkk som intogs ht 1967 och ht 1968 (s:a 2 064), ur vilken sedan fpp utvaldes för experimentet — ett förfaringssätt som ej kan tillämpas med hänsyn till lärarutbildningens organisation.

För att minska variationsmöjligheterna med hänsyn till antalet faktiskt tillgängliga studerande och experimentets genomförbarhet och för att på så sätt garantera tillfredsställande analysmöjligheter (t ex styrka i F-testet) stratifierades den tillgängliga populationen med avseende på kön och utbildningsbakgrund vid inträdet. Populationen indelades i tre strata: (1) kvinnliga studerande med slutbetyg från gymnasiet och (2) manliga studerande med slutbetyg från gymnasiet (denna grupp motsvarar SÖ behörighetsgrupp b och e) samt (3) studerande med slutbetyg från en fack- eller folkhögskola eller motsvarande utbildning (SÖ behörighetsgrupp a, c och d).

Eftersom kvinnliga studerande utgör den absoluta majoriteten på denna utbildningslinje och eftersom vi ville undvika eventuella variationer på grund av kön eller utbildningsbakgrund, uteslöts strata 2 och 3 ur samplingen.

Antalet studerande i stratum 1, den kvinnliga gruppen med slutbetyg från gymnasiet, var för år 1967 n=72 och för 1968 n=61. Ur denna grupp av sammanlagt n=133 drogs ett stickprov av n=96 som deltog som fpp i experimentet. Experimentet genomfördes för respektive terminskurs i början av andra terminen vid LHM. Eftersom experimentets omfattning medfört en uppdelning i två faser, slumpades fpp som deltog i fas 1 (vt 1969) via en slumptabell till grupp 2 (handledning) och 3 (självkonfrontation via ITV/VR). De fpp som deltog i fas 2 (vt 1970) slumpades till grupp 1 (handledning *och* självkonfrontation via ITV/VR) och 4 (ingen påverkan). Båda terminskurserna är alltså representerade i experimentets båda oberoende variabler. Detta förfaringssätt åskådliggörs i figur 1.

4.2 Utbildningsbakgrund

Likvärdigheten mellan experimentets olika grupper granskades med hänsyn till (1) ålder, (2) utbildning vid intagning till klasslärarlinjen (intagningspoäng) och (3) socialgruppstillhörighet. I tabell 1 ges medeltalen för lkk ålder i respektive grupp.

Tabell 1. Åldersfördelning i medeltal (år)

År	Grupper 1	2	3	4
1969		21.46	20.67	
1970	21.07			20.91

Någon påvisbar skillnad finns inte mellan grupperna i fråga om åldern ($\chi^2 = .001$).

Intagningen till lärarutbildningen sker som nämnts centralt via Skolöverstyrelsen. Som sammanfattande uttryck i meritvärderingen får varje sökande en intagningspoäng.

Intagningspoängen har i experimentet använts för en granskning huruvida de lkk, som ingår i de olika grupperna, skiljer sig med hänsyn till initialförutsättningarna.

Tabell 2. ANOVA för intagningspoäng

Variations-orsak	df	SS	MS	F
Mellan	3	.16	.05	.56
Inom	92	7.93	.09	
Total	95	8.08		

För förkortningar, se bilaga 1.

Som framgår ur tabell 2 kan grupperna med hänsyn till intagningspoängen betraktas som likvärdiga.

Lkk socialgruppstillhörighet kontrollerades med hjälp av data för målsmannen. För kodning av lkk socialgruppstillhörighet har Carlssons yrkeskategorisystem använts (Carlsson, 1958, ss 49—51). Socialgruppsfrekvensen ges i tabell 3.

Tabell 3. Lärarkandidaternas socialgruppsfrekvens med hänsyn till målsmannens socialgruppstillhörighet

Experiment-grupp	Social-grupp I	Social-grupp II	Social-grupp III
1	9	12	3
2	12	5	7
3	12	10	2
4	9	6	9

För eventuella skillnader testades fördelningen med χ^2-testet på 5 %-nivån. Någon skillnad mellan experimentets olika grupper med hänsyn till socialgruppstillhörighet finns ej eftersom χ^2 (6 df) = 11.21.

De lkk som deltog i experimentet har vid LHM undervisats enligt de studieplaner som gäller för lärarutbildningen fr o m höstterminen 1967. Detta innebär att det ej har förekommit övningsundervisning före experimentets början.

Den praktiska utbildningen under första terminen på mellanstadielärarlinjen (M 1), består av ca 6 tim auskultationer på mellanstadium. Den metodiska undervisningen under M 1 samordnas med pedagogikundervisningen på så sätt som beskrivs i studieplaner för klasslärarutbildningen, SÖ, Stockholm, 1967 på sidan 421. Samordningen gäller både innehållet i undervisningen och uppgifterna under auskultationerna. Någon metodikundervisning som direkt har till syfte att förbereda lektioner har inte förekommit.

Auskultationsuppgifterna avser observationer av lärare och elever i klassrummet. Efter varje auskultation redovisar lk sina iakttagelser av det sätt på vilket undervisningen genomförs. Dessa redovisningar diskuteras bl a i form av jämförelser. Man diskuterar de läromedel, framför allt läroböcker och bredvidläsningslitteratur, som den studerande har sett i användning. I diskussionsform görs jämförelser angående innehåll, bildmaterial och arbetsuppgifter. Dessa sammanfattande diskussioner äger rum sedan auskultationerna har genomförts.

Under terminskurs M 2 har före experimentets början ej någon metodikundervisning förekommit.

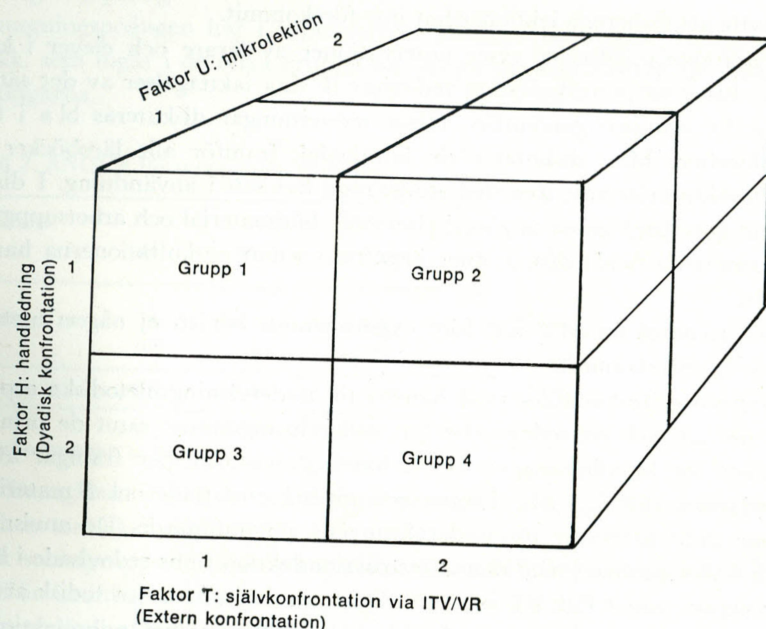
Experimentet förbereddes med hänsyn till undervisningsmetodiska aspekter genom att lkk fick en redogörelse för undervisningsämnet samt de principer som gäller för handledningen. Dessa handledningsprinciper framgår av SÖ Handledarbok (1967, s 34). Denna genomgång omfattade också materialanvisningar samt litteratur för undervisningens genomförande. De anvisningar som lkk fick i samband med denna instruktionslektion finns redovisade i kap 5. Under experimentet fick lkk inga ytterligare anvisningar av metodisk art.

Det samtal som handledaren hade med lkk före själva mikrolektionerna rörde enbart undervisningsmaterial, elevernas placering i studion, elevnärvaro samt kontroll att lkk hade uppfattat inspelningstiden, dvs när inspelningen påbörjades och avslutades enligt skriftliga anvisningar. Samtalet om undervisningsmaterialet avsåg endast att hjälpa lkk att finna planscher m m.

Att lkk i så liten utsträckning har kommit i kontakt med en mera systematisk övning av olika undervisningsfärdigheter inom lärarutbildningen, men dock haft tillfälle att acklimatisera sig på lärarhögskolan, har varit skälet till att lkk vid andra terminens början valdes som fpp för experimentet. Detta stickprov skulle kunna ge en uppfattning om lkk inträdeskunskaper (initialkunskaper) angående observation och analys av undervisningsprocesser och -beteenden.

5. Undersökningens experimentella utformning

I syfte att studera undersökningens problemställning under experimentella betingelser utformades uppläggningen som åskådliggörs i figur 2.



Figur 2. 2×2×2-faktoriell design med repeterad mätning i den sista faktorn

Designens tre faktorer är:

Faktor H: traditionell handledning, där
h_1: handledning,
h_2: ingen handledning.

Faktor T: externt förmedlad självkonfrontation via ITV/VR, där
t_1: självkonfrontation,
t_2: ingen självkonfrontation.

Faktor U: mikrolektioner (tidslängd 15 min), där
u_1: mikrolektion 1,
u_2: mikrolektion 2.

Faktorerna T och H är fixa och manipulativa variabler, och designen är med hänsyn till dessa faktorer fullt kombinerad och fullständigt balanserad. Den tredje faktorn (faktor U) är fix. Med hänsyn till experimentet betyder detta att det inom varje påverkningskombination finns 24 oberoende observationer som görs under experimentellt identiska betingelser. Faktor U betyder att varje lk håller två mikrolektioner. Designen medför däremot att individfaktorn är kopplad under faktorerna H och T, vilket i sin tur medför att det inte finns interaktionstermer mellan dessa faktorer och individfaktorn. Modellens implicita antagande är samma homogenitet i kovariansmatriserna som är relaterade till den repeterade mätningen (Winer, 1962, s 338). Att testa den experimentella datauppsättningen på lika varians i cellerna har i allmänhet en relativt begränsad praktisk användbarhet. Enligt senare diskussion (Box, 1953, 1954) borde en variansanalytisk bearbetning i varje fall utföras, dvs också utan föregående homogenitetstest och under alla förhållanden i sådana fall där antalet individer i cellerna är lika (se Hays, 1970, s 318). Om dessutom experimentet medför ett relativt stort antal frihetsgrader (df > 30) för skattningen av variansen, som hänför sig till de enskilda interaktionerna mellan påverkan och individer, behövs det i allmänhet ingen speciell granskning av F-testets konstruktion. Slutligen är F-testet för denna uppläggning (alla faktorer fixa) relativt robust, dvs okänsligt mot avvikelser från homogenitetskravet och kravet på normalfördelning. I ett experiment, upplagt för att undersöka en kombinerad effekt av två eller flera experimentella variabelkombinationer, bör man uppmärksamma om individerna utsättes för fler än *en* variabelkombination, eftersom användbarheten av en design med repeterad mätning får en viss begränsning genom eventuella överföringseffekter, som lätt kan sammanblandas med påverkningseffekterna. Men detta gäller framför allt inlärningsexperiment. Huvudsyftet med den repeterade mätningen i den aktuella designen är att kontrollera individuella skillnader mellan de olika experimentella enheterna (fpp). I beteendevetenskapliga undersökningar är skillnaderna mellan dessa enheter ofta relativt stora jämfört med de experimentella effekter som man önskar påvisa (Winer, 1962, s 300).

Den i figur 2 presenterade designen måste betraktas som en "cross-over design" eftersom det ej kan uteslutas att det finns vissa additiva överföringseffekter. För att öka designens precision och för att kunna lösa vissa mättekniska problem utökades den ursprungliga designen med ytterligare två faktorer. Huvudformuläret (F III) i experimentet består av 79 påståenden. Mätinstrumentet ingår i designen som faktor V med 79 nivåer. Varje påstående innehåller två olika aspekter; (a) en för förekomst av eller kvalitet hos ett attribut och (b) en för värdering av detta attribut. Dessa två aspekter representeras av faktor A med två nivåer. Hela ANOVA-modellen kan mot denna bakgrund skrivas om A, U, T, H, I(TH), V, där I betecknar individfaktorn.

Analysen möjliggör genom designens utökning också en kontroll av mätinstrumentens effekter, där framför allt interaktionseffekterna är av intresse. Signifikanta huvudeffekter i faktorerna A och V samt den enkla interaktionseffekten VA borde vara självklara. En sammanfattning av experimentets variansanalytiska uppläggning presenteras i tabell 4. Utformningen av faktorerna H, T och U beskrivs mera ingående nedan, medan det för beskrivningen av faktorerna I respektive A och V hänvisas till kapitel 4 respektive 8.

Tabell 4. Experimentets variansanalytiska uppläggning

Index	A	U	T	H	I	V
Antal nivåer	2	2	2	2	24	79
Populationsstorlek	2	2	2	2	∞	79

5.1. Beskrivning av faktorerna H, T och U

5.1.1 Faktor H: Traditionell handledning

Innan lkk börjar övningsundervisningsperioden har metodikläraren, som är handledare under denna tid, ett seminarium för samtliga 24 lkk i kandidatgruppen.

Under seminariet behandlas allmänna undervisningsprinciper, bl a lektionsförberedelse, val av undervisningsämnen, stimulering och arbetsuppgifter. Seminariet fungerar som en sammanfattning av vad som tidigare behandlats i metodikundervisningen. På denna genomgång bygger lkk övningsundervisningen. Anvisningar för genomförandet av handledningen finns i Skolöverstyrelsens Handbok för handledare vid praktisk lärarutbildning. Klass- och ämneslärarlinjerna (Stockholm: SÖ-förlaget, 1967).

De konkreta uppgifterna för förberedelse och genomförande av övningsundervisningen får lkk från handledaren senare på respektive praktikskola. Vid genomförandet av övningsundervisningen finns handledaren till hands. Han observerar lkk i klassrummet och gör under dessa observationer anteckningar för senare lektionsgenomgång. De aspekter som enligt handledarboken (s 32) bör uppmärksammas vid mera lärarstyrd undervisning presenteras i ruta 1.

Undervisningsfrågor av mera principiell natur som kan vara av intresse för alla lkk behandlas i regel i seminarier som ansluter sig till denna övningsperiod.

5.1.1.1 *Handledning under experimentet*

Finns det någon likhet mellan handledning under lkk övningsundervisning och handledning som experimentell påverkan? De skriftliga anvisningarna som handledaren och lkk fick i samband med den första konferensen (en vecka före undersökningens början) redovisas nedan. Denna konferens kan sägas motsvara seminariet som lkk har tillsammans med metodikläraren före övningsundervisningsperiodens början. Under konferensen informerades lkk om nedan redovisade metodiska anvisningar, handledningsprinciper och materialanvisningar.

Ruta 1. Iakttagelser vid mera lärarstyrd undervisning

Formella aspekter och läromedels användning, t ex lärarens
tidsdisposition,
röst och språk,
frågor och svar för att få god elevaktivitet,
användning av tavlan,
användning av lärobok och annan litteratur,
experiment och demonstrationer,
utnyttjande av läromedel.

Innehållet i undervisningen, t ex
ämnets samverkan med andra ämnen,
översikt och sammanhang,
fördjupningsmoment,
lärarens exemplifiering.

Ämnesmetodiska problem

Instruktionernas utformning

Arbetsuppgifternas utformning

Hur läraren aktiverar eleverna, t ex
genom frågor och samtal,
genom problemställande,
genom arbetsuppgifter.

Hur samspelet med eleverna utformas, t ex
hur elevernas motivation utnyttjas,
hur beröm och klander användes,
hur självförtroende skapas.

Hur läraren skapar omväxling

Hur arbetet organiseras

Elevernas reaktioner på undervisning, t ex
frimodighet,
arbetslust,
intresse,
frågor,
spontanitet.

Kontroll och värdering av t ex
kunskaper och färdigheter,
poängsättning och betyg.

Lärarens behandling av oväntade situationer

Hemuppgifternas
omfattning,
individualisering.

På samma sätt som under övningsundervisningen observerade handledaren lkk under experimentet, dvs gjorde de anteckningar, som ansågs nödvändiga för den efterföljande handledningssessionen.

Den tid som handledaren hade till sitt förfogande för att gå igenom lektionen med lk kan anses lika lång som den som vanligen står till förfogande under övningsundervisningen. Handledning som experimentell betingelse genomfördes alltså formellt inom samma ram som den *traditionella* handledningen under övningsundervisningen.

5.1.1.2 *Beskrivning av handledningsprocessen*

Undervisningsämnena under experimentet omfattade Norrlands djur. Som konkretionsmaterial kom framför allt planscher och kartor samt uppstoppade djur till användning.

En klassuppsättning läroböcker (Linnman, N. & Wennerberg, B. Grundskolans naturkunskap, årskurs 4—6. Stockholm: Bergvall, 1962) och planscher fanns under experimentets genomförande i ITV-studion.

I de fall där lkk utnyttjade möjligheten att ge arbetsuppgifter till eleverna, samlades dessa in av handledaren för en eventuell genomgång under den anslutande handledningsperioden. Enligt handledarens bedömning undervisade alla lkk under experimentet på ett mycket likartat sätt. (Samma fenomen poängterades f ö upprepade gånger av en för experimentet sammanställd expertpanel, kap 7.2.) Lektionerna genomfördes huvudsakligen som samtalslektioner.

5.1.1.3 *Anvisningar för handledaren*

Under experiment- och kontrollgruppens gemensamma lektion före försöket informerar handledaren om

I. handledning under övningsundervisning (Vad observerar handledaren?)

II. undervisningsämnet under försöket

III. lk förberedelse inför undervisningen

IV. lk genomförande av undervisningen

1. lk informeras om de iakttagelsekategorier som *Handledarboken* anger för *lärarstyrd undervisning:* Formella aspekter och läromedelsanvändning, innehållet i undervisningen, ämnesmetodiska problem, instruktionernas utformning, arbetsuppgifternas utformning, hur läraren aktiverar eleverna, hur samspelet med eleverna utformas, hur läraren skapar omväxling, hur arbetet organiseras, elevernas reaktioner på undervisning, kontroll och värdering, lärarens behandling av oväntade situationer samt hemuppgifter. Det framhålles för lk att handledaren gör en bedömning av vad som vid det aktuella tillfället tas upp till behandling, samt att lk egna synpunkter på sin undervisning är av stor betydelse för handledningen. TV-studions speciella inverkan på undervisningen behandlas ej.

2. Ämnet är naturkunskap och omfattar Norrlands djur (enligt särskild förteckning). Lk undervisar eleverna om ett djur. Vilket djur som skall behandlas framgår av en fördelning uppgjord av försöksledaren. Detta arrangemang måste tillgripas för att förhindra att lk, som undervisar samma klass, väljer samma djur. Undervisningsmaterial som lk önskar använda lånas på Lärarhögskolans försöks- och demonstrationsskola.
Eleverna har inte arbetat med detta lärostoff i åk 4.

5.1.1.4 *Anvisningar för lärarkandidater*

"Vid undervisningens början kommer eleverna att befinna sig utanför TV-studion. Din lektion omfattar 20 minuter. Av denna tid använder Du de första fem minuterna till att bekanta Dig med eleverna och under de återstående 15 minuterna genomför Du lektionen. Det är denna del som spelas in.

Förberedelse:

Du planerar Din undervisning som katederlektion (klassundervisning). Gör klart för Dig vad Du vill lära eleverna och välj ut material som Du önskar använda.
När Du planerar kan Du fundera på följande:

I vilken miljö lever djuret?
Hur anpassar sig djuret till miljön (föda, bo, utseende)?
Hur "sköter" djuret sina ungar?
Har djuret några fiender?

I vilken ordning och hur Du skall presentera dessa moment bestämmer Du själv. Du skall inte känna Dig tvingad att behandla alla momenten.

Genomförande:

När Du bekantar Dig med eleverna, kan Du presentera Dig för dem samt tala med dem om *TV-studions utrustning*. Låt eleverna se sig omkring i studion. När *GRÖN LAMPA* lyser börjar Du undervisningen och genomför den så som Du planerat.
När *RÖD LAMPA* lyser pågår inspelning.
När *GUL LAMPA* lyser har Du två minuter på Dig att avsluta undervisningen.
OBS! GRÖN LAMPA lyser – inspelning börjar
GUL LAMPA lyser – inspelningen upphör om två minuter.

För att undvika tidsförskjutningar vid inspelningarna bör Du infinna Dig vid TV-studion i *god tid* (30 min före lektionens början)."

Under första lektionen noterades olika moment angående inledningsformer, t ex: Hur inledde lk lektionen? Var det en intressant inledning, som lämpade sig för att öka elevernas uppmärksamhet? På vilket sätt följde lk upp denna inledning? På vilket sätt utnyttjades planscher, tavlan, läroböcker, arbetsmaterial etc?

Handledaren gjorde under lektionen en skiss av elevernas placering. Mycket aktiva och mycket passiva elever markerades och exempel på aktivitet och passivitet antecknades som underlag till den efterföljande handledningsdiskussionen. Att börja diskussionen med utgångspunkt från denna skiss visade sig snart (enligt hl anteckningar) vara ett mycket lämpligt sätt att angripa kontaktproblemet utan att lkk behövde få en negativ upplevelse, dvs upplevelsen att de har gjort något fel.

Under handledningen i anslutning till den första mikrolektionen behandlades dessutom uppmärksamhetsprincipen, dvs: Har lk under lektionen lyckats med att fånga elevernas uppmärksamhet? Vad fångade elevernas uppmärksamhet? Borde lk eventuellt ytterligare ha ökat elevernas uppmärksamhet?

Handledningsdiskussionen inleddes med att handledaren uppmuntrade lk att berätta sina egna erfarenheter. Följande problemställningar bildade bl a utgångspunkt till handledarens inledande frågor: Hade lk mera detaljerat observerat de aktiva och passiva eleverna respektive över huvud taget märkt de aktiva och passiva eleverna? Vilka åtgärder kom till användning för att t ex dämpa överaktiva elever och stimulera passiva? På samma sätt behandlades användningen av tavlan och konkretionsmaterialet. Under diskussionen uppmuntrades lk till att ge egna förslag till förbättringar. Det förekom dock mycket sällan att lk gav egna förslag till förbättringar. I stället fanns det en utpräglad tendens till att hellre vilja berätta om upplevelsen av undervisningsprocessen med hänsyn till svårigheter att komma ihåg vad som skulle sägas, eller strävan att fråga *alla* elever samt upplevelse av tidsproblemet, dvs att ej glömma bort inspelningstiderna. Handledningen i anslutning till andra mikrolektionen omfattade utöver det som anfördes i samband med den första lektionen: (1) elevernas kunskaper, dvs på vilket sätt lk integrerade elevernas kunskaper i undervisningen och (2) olika aktivitetsformer. Varje handledningstillfälle avslutades med några råd. Exempel på sådana råd efter första lektionen är:

1. En mera fartfylld inledning för att rycka med eleverna.
2. Referera till aktuella händelser, t ex TV-serien om Norrlands djur som just under experimentet sändes av Sveriges Radio.
3. Variera aktiviteterna, låt eleverna gå fram och skriva på tavlan, visa och peka på planscher.

Samma lk fick i samband med andra lektionens handledningstillfälle följande råd:

1. Något fastare ledning av lektionen så att den styr mot det mål som sattes upp för den.
2. Försök att rycka med de passiva eleverna genom att ställa frågor till dessa elever, skicka fram dessa elever till tavlan och planscherna.
3. Försök att tänka igenom varför eleverna är passiva respektive överaktiva.

Det bör nämnas att den enkla skissen av elevernas lokalisering med hänsyn till de aktiva och passiva eleverna visade sig vara ett mycket centralt moment i handledningen. Samtliga lkk lektioner visade sig vara så pass lika varandra att handledningen av alla lkk kunde genomföras i enlighet med den ovan beskrivna rutinen. I regel blev alltså följande undervisningsprinciper föremål

för handledningsdiskussionen: (1) lektionsinledning, (2) uppmärksamhet, (3) elevernas kunskaper, (4) elevernas aktivitet och (5) kontakten mellan lärare och elev. Detta bör dock ej tolkas så att handledaren alltid började med en och samma undervisningsprincip. Strävan har varit att försöka bedöma vilket moment som lämpade sig bäst som inledning.

Slutligen kan också anföras att denna handledning, baserad på den ovan beskrivna rutinen som är konkretiserad i handledarens undervisningsprotokoll, kan betraktas som ett lämpligt hjälpmedel för att minska felvariationer som bottnar i de förändringar som handledarens bedömningar är utsatta för.

5.1.2 Faktor T: Externt förmedlad självkonfrontation via ITV/VR

Med externt förmedlad självkonfrontation avses här konfrontation med det egna beteendet i undervisningssituationer via intern television och videobandning. Processen innebär en *extern* konfrontation med det egna expressiva beteendet, som individen har lärt sig att "inte se" (Holzman, 1969, s 207). Konfrontationserfarenheter skulle kunna beskrivas som en de-automatiseringsprocess av det vanliga sättet att uppleva sig själv. Resultatet av en extern redirigering av uppmärksamheten, förmedlad via ITV/VR, måste åtminstone temporärt betraktas som en desorganisation av olika beteendefunktioner. Under de-automatiseringsprocessen kan då nya funktioner byggas upp. Denna process skapar således tillfällen till förbättring, återtagande av eller kvarhållande av status quo för en eller flera funktioner. Vid självkonfrontationen föreligger en situation, där de-automatiseringsprocesser antagligen har uppstått på grund av bristande balans mellan individens faktiska beteende och hans automatiserade själv-image.

Den teknik ("självkonfrontation via ITV/VR"), som används i experimentet skiljer sig från en "simultan självspegling" på följande sätt:

1. Individen har ej längre tillfälle att successivt redigera den inkommande auditiva informationen, dvs rösten, och den visuella informationen, dvs kroppens rörelseschema, fysionomi, gester samt andra expressiva kvaliteter.
2. Att bli konfronterad med ljud- och videobandade stimuluskonfigurationer skapar för individen en helt ny referensram ("context") för perception och värdering av de audiovisuella stimuluskonfigurationerna. Rösten får permanens och på samma sätt som vid betraktande av ett fotografi av sig själv tvingas personens uppmärksamhet in på sådana aspekter som han i annat fall (spegling) hade modifierat eller ignorerat.
3. Videobandningens möjligheter att överföra det förflutna till nuet är en alldeles ny erfarenhet, som också på något sätt torde påverka individens erfarenheter.

Denna T-faktor kan alltså sägas innebära en extern självdistansering i rum och tid (för beskrivning av den tekniska anordningen, se kap 6).

5.1.3 Faktor U: Mikrolektioner

5.1.3.1 *Mikrolektionernas funktion i lärarutbildningen*

Begreppet "micro-teaching" (mikrolektioner) utvecklades ursprungligen (år 1963) av Allen & Bush i samband med *The Stanford Teacher Education Program* vid Stanford University (Allen & Ryan, 1969, s 4).

Mikrolektionstekniken är en nedbantningsteknik och enligt Allen & Ryan (1969, ss 33—59) definierad genom följande grundelement:

1. *Lektion*
   Mikrolektioner måste ses som riktiga lektioner även om undervisningssituationen i olika avseenden kan ha begränsats, t ex i förväg bestämda samarbetsformer mellan studerande och elever.
2. *Komplexitet*
   Mikrolektioner minskar komplexiteten i förhållande till vanliga lektioner, dvs antalet elever, undervisningämnets omfattning och lektionens längd reduceras.
3. *Fokusering*
   Mikrolektioner avser en träning av undervisningsfärdigheter, undervisningsstrategier eller demonstration av undervisningsmetoder.
4. *Kontroll*
   Mikrolektionstekniken innebär en ökad kontroll av övningsundervisningen. Variabler som t ex tidsintervaller, elever och olika feedbackmetoder kan kontrolleras.

Tekniken har tillämpats i allt högre grad såväl i lärarutbildningen som i utbildningen av psykoterapeuter och medicinare. Elevgruppen för lkk övningsundervisning i den vanliga lärarutbildningen är den traditionella skolklassen med ca 25—30 elever. Begreppet mikrolektion eller mikroundervisning är ej något statiskt begrepp. Det används för beskrivning av många olika nedbantningsformer av övningsundervisning ägnade att utveckla lärarens undervisningsskicklighet. Begreppet har gett upphov till förhoppningen att kunna överbrygga klyftan mellan den mera teoretiska och den praktiska utbildningen av lkk.

För att bättre förstå värdet av mikrolektionstekniken kan det vara angeläget att kort rekapitulera några aspekter av den traditionella lärarutbildningen. Den bygger på (1) teoretiska kurser och (2) övningsundervisning samt (3) observation och auskultationer. Pedagogiken har i lärarutbildningen bl a till uppgift

1. att ge den blivande läraren vetenskapligt grundade utgångspunkter för sina åtgärder som uppfostrare och undervisare,
2. att ge kunskaper och färdigheter som möjliggör en tillämpning av pedagogikens principer på praktiska undervisningsproblem,
3. att ge kunskaper och färdigheter som möjliggör en analys och utvärdering av effekten av utbildningsprocessen,
4. att ge orientering om det pedagogiska forsknings- och utvecklingsarbetet,
5. att befrämja personlig utveckling och självkännedom så att den blivande läraren utifrån sina personliga förutsättningar och under eget ansvar kan handha fostran och undervisning.

Som framgår ur dessa fem punkter (se Klasslärarutbildning, 1968, s 393) ligger tonvikten i lärarutbildningen, när det gäller pedagogik, på ämnets tilllämpning. Det framhävs att den blivande läraren bör tillägna sig "färdigheter" i de ovan angivna aktiviteterna.

Avgörandet om och i vilken utsträckning ett visst förfaringssätt i undervisningen kan betraktas som önskvärt och effektivt, tycks allt mera bli styrt av i vad mån lkk kan förverkliga ett metodiskt förfaringssätt och allt mindre av lkk prestationer i t ex skriftliga tentamina (se Lee, 1971). För att uppfylla målsättningen, borde lkk utbildning således i mycket högre grad än hittills interfolieras med sådana moment som möjliggör att lkk t ex kan "experimentera" med sig själva i olika lärarroller. En sådan strukturförändring skulle innebära, att man på ett mera adekvat sätt uppmärksammade variationer i lkk individuella behov. En sådan förändring av tonvikten skulle också betyda att man på ett välgörande sätt avlägsnade sig allt mera från "standardkurser", "standardläroböcker" och "standardexaminationer" utformade för en mytisk "standardstuderande" (för diskussion se Lee, 1971, ss 18—20). Koncentreras utbildningen på en utveckling av lk som "personlighet" i stället för en utbildning av lk som en "lärling inom ett hantverk", krävs det en systematisk integration av pedagogisk teori och praxis, dvs (1) systematiska observationer, (2) övningsundervisning av mindre komplex natur och (3) en styrd bearbetning av utbildningens praktiska moment (för diskussion se Bjerstedt, 1970 b, ss 222—223).

Att kunna laborera med olika undervisningsformer är ej så lätt i den traditionella utbildningen. Som Cohen, L. (1969, s 48) visade, är varken klassläraren eller praktikskolan förtjust i att en lk kanske försöker pröva nya metoder eftersom båda är ansvariga för kontinuiteten i elevernas utbildning. I detta sammanhang syns mikrolektionen kunna fylla en mycket viktig funktion. Att observera undervisningsprocessen under laboratoriemässiga betingelser bör kunna medföra en värdefull förbättring av lärarutbildningen. Därvid kan nämligen den studerande utforma sina studier av undervisningsprocessen enligt samma metodik som forskaren använder dvs:

1. att utveckla en arbetshypotes

2. att iaktta och protokollera undervisningsbeteendet i den pågående undervisningsprocessen
3. att analysera dessa protokollerade data och
4. att dra slusatser av dessa data.

Framför allt den pedagogisk-psykologiska forskningen kan med fördel använda denna laboratorieteknik för att få den komplexa undervisningsprocessen under experimentell kontroll. Speciellt som forskningsteknik syns denna metod ge upphov till många olika variabelkombinationer. Hela mikrolektionsstrukturen kan manipuleras för att med hjälp av experimentella designer försöka besvara problemställningar som t ex: Vad betyder olika feedbackformer för lkk modifikation av referensram och beteendemönster?

5.1.3.2 *Elevgrupper*

Mikrolektionstekniken skall förmedla äkta undervisningserfarenheter. Med denna målsättning blir eleverna ett väsentligt moment. Eleverna skall garantera att undervisningen blir en realistisk undervisningserfarenhet. Speciell uppmärksamhet bör ägnas åt urvalet av de elever som skall bilda en undervisningsgrupp. Vägledande för antal och val av eleverna till experimentet blev följande synpunkter:

1. Elever som deltar i mikrolektionerna bör vara representativa för det stadium den blivande läraren skall undervisa på.
2. Även årskursen, från vilken eleverna kommer, är en relevant faktor. En fjärde årskurs valdes dels med hänsyn till förförsöket, dels med hänsyn till att elever i årskurs 4 i allmänhet betraktas som harmoniska och lätt anpassbara till ITV-studion.
3. Elever som deltar i experimentet bör tidigare ha stiftat bekantskap med auskulterande lkk och bör ha deltagit i demonstrationslektioner i LHM klassrum som är utrustade med one-way-screen.
4. Eleverna som deltar bör ej få någon speciell uppgift eller anvisning från experimentledaren. Om mikrolektionen skall förmedla lkk upplevelser av en vanlig undervisningssituation är det viktigt att eleverna ej tränas för ett bestämt reaktionsmönster. Intentionen är att eleverna reagerar (ger feedback) i enlighet med den egna kognitionen av lkk beteende.
5. Eleverna bör delta inom ramen för den ordinarie undervisningen.

De elevgrupper (halvklasser) som år 1969/70 deltog som undervisningsgrupper kom samtliga från Munkhätteskolans 4:e årskurs. Varje årskurs består av 6 parallella klassavdelningar med ca 25—30 elever per avdelning. Sammanlagt deltog under de båda läsåren ca 360 elever. Varje klass delades slumpmässigt upp i två halvklasser.

En slumpmässig generering av elevgrupper över hela årskursen diskuterades som ett alternativt förfaringssätt, men visade sig av flera skäl vara mindre önskvärd. Dels skulle svårlösta schematekniska problem ha uppstått, dels skulle en sådan teknik ha medfört en onaturlig situation för eleverna.

Urvalet genomfördes med hjälp av klasslistornas alfabetiska ordning. Elevernas likvärdighet avseende undervisningsprestation undersöktes. Som utgångsmaterial valdes elevernas betyg. Betygsfördelningen testades med ANOVA.

Tabell 5. ANOVA för elevernas betygsfördelning år 1969 och 1970

Variations-orsak	df	SS	MS	F
Mellan	3	.01	.003	.016
Inom	20	.35	.18	
Total	23	.36		

Mellan: Elevgrupp 1 och 2 år 1969 och elevgrupp 1 och 2 år 1970, dvs 4 elevgrupper.
Inom: De sex parallella klassavdelningarna (a–f).

Någon skillnad mellan elevgrupperna avseende betyg, fördelade på experimentets fyra grupper, finns inte.

Elevernas socialgruppstillhörighet kodades enligt samma kategoriseringssystem som har använts för lkk. Fördelningen redovisas i tabell 6. För att kunna testa denna blev en sammanslagning av socialgrupp 1 och 2 nödvändig.

Tabell 6. Elevernas socialgruppstillhörighet

Klass (Elevgrupp)	Socialgrupp 1968 1, 2	3	1970 1, 2	3
4 a I	6	9	6	10
4 a II	5	9	6	9
4 b I	9	5	5	10
4 b II	4	9	2	11
4 c I	6	8	6	7
4 c II	8	6	7	6
4 d I	7	6	5	7
4 d II	7	6	5	7
4 e I	3	11	4	9
4 e II	7	7	7	5
4 f I	8	6	5	6
4 f II	10	3	4	7

Fördelningen testades med χ^2-testet på 5 %-nivån. Någon skillnad avseende socialgruppstillhörighet mellan elevgrupperna i de fyra experimentella betingelserna kan inte konstateras, eftersom χ^2 (30 df) = 20.11.

För att eventuellt möjliggöra vissa jämförelser med hänsyn till lkk agerande inför en viss elevgrupp bestämdes att 8 lkk, 4 ur respektive påverkningsgrupp, skulle undervisa de elever som tillhörde en och samma klassavdelning. Som framgått indelades årskursen i 12 elevgrupper (halvklass). Till varje halvklasspar slumpades en klassavdelning samt de lkk som skulle undervisa respektive elevgrupp. Fördelningen anges i tabell 7.

Tabell 7. Uttagning av klassavdelningar och lärarkandidater till elevgrupper

Elev-grupp	Klassav-delning	lkk nr i grupp 1 och 3	2 och 4
1	c	1 och 13	25 och 27
2	c	2 och 14	26 och 38
3	d	3 och 15	27 och 39
4	d	4 och 16	28 och 40
5	f	5 och 17	29 och 41
6	f	6 och 18	32 och 42
7	a	7 och 19	31 och 43
8	a	8 och 20	32 och 44
9	b	9 och 21	33 och 45
10	b	10 och 22	34 och 46
11	e	11 och 23	35 och 47
12	e	12 och 24	36 och 48

Denna fördelning medger en granskning av eventuella effekter med hänsyn till televisions- och handledningsfaktorn inom klassens ram.

Eleverna verkade överlag uppleva deltagandet som mycket positivt. Vanligen fanns eleverna på plats redan ca 30 minuter före fastställd lektionstid. Varje lk undervisade en och samma elevgrupp vid båda undervisningstillfällena för att lk lättare skulle kunna percipiera och värdera undervisningseffekterna och effekten av sitt eget beteende. För att undvika att en elevgrupp blev undervisad två eller flera gånger i ett och samma ämne blev det däremot nödvändigt att fördela ämnena systematiskt med hänsyn till elevgrupperna.

5.1.3.3 *Undervisningsämne*

Som ämnesområde valdes för experimentet biologi och som undervisningsämne Norrlands djur: fjäll-lämmel (I), björn (II), varg (III), lo (IV), ren (V), kungsörn (VI), ripa (VII) och järv (VIII). (De inom parentes angivna romerska siffrorna refererar till tabell 8.) Dessa ämnen valdes eftersom metodiken av lektorer bedömdes som förhållandevis enkel. Undervisningstekniken är med hänsyn till de olika djurarterna densamma. Även systematiken för undervisningens uppläggning bedömdes som likartad med hänsyn till lkk utbildningsstadium.

Ämnet biologi, speciellt djur betraktas som ett engagerande ämne för elever i årskurs 4. Detta ämnesval kan anses vara mognads- och motivationsmässigt väl avpassat för stadiet. Biologi har dessutom fördelen att alla lkk har tillgång till ett rikligt och mycket likartat konkretionsmaterial. Dessutom är sådant material lätt tillgängligt på LHM. Även eventuellt nödvändig litteratur för snabb orientering finns i tillräckligt stor mängd på LHM bibliotek. Å andra sidan kan ämnesramen ses som mycket flexibel. Exempel på undervisningsmoment skulle kunna vara: (a) utseende — rörelse — flykt — miljö, (b) mat — kampen för tillvaron — fiender, (c) bo — övervintring, (d) ungar — matning. Alla fyra punkterna kan naturligtvis behandlas, men lkk fick själva bestämma tyngdpunkten i undervisningen. Enda villkoret för undervisningsförberedelse var att lektionen skulle läggas upp som en "mera lärarstyrd" lektion. Det poängterades dock att mindre gruppaktiviteter skulle kunna placeras in i lektionen.

Lkk skulle med denna utgångssituation ha möjlighet att agera i olika situationer, nämligen: aktivera elever, utforma frågeställningsteknik, visa en svarsbehandlingsteknik, lära eleverna att se, iakttaga och dra slutsatser. Lkk skulle kunna aktivera elevernas förkunskaper, fokusera elevernas uppmärksamhet samt leda eleverna till att förstå mera komplexa problemställningar. Elever i denna ålder skulle kunna bidra med många egna uppslag till undervisningen, dels på grund av att de har haft egna upplevelser av eller har hört talas om djur, dels eftersom de utformar en viss fantasi kring djuren. Jämfört med andra ämnen som t ex historiska gestalter eller religionskunskap skulle man kunna förvänta sig mycket aktiva elevgrupper.

5.1.3.4 *Lektionslängd*

Det stora flertalet lkk som ingick i experimentet hade ej undervisat tidigare och kände ej eleverna. För att bekanta sig med eleverna fick lkk 5 min uppvärmningstid innan den effektiva, videobandade undervisningstiden började, begränsad till 15 min. I många olika sammanhang har man ifrågasatt om det överhuvudtaget är möjligt att undervisa inom en period av 5—20 min. Kanske beror detta på att man vanligen tänker sig att kunskaper kan förmedlas enbart i 40 eller 50 min perioder. Olika erfarenheter i samband med undersökningar och experiment har dock visat att tidsvariabeln är av underordnad betydelse (Allen & Ryan, 1969, s 40). Troligen är det så att vi tillmäter tidsvariabeln alltför stor betydelse.

Tabell 8. Fördelning av elevgrupper, lärarkandidater och ämnen för mikrolektionerna

	Måndag					Tisdag				
Tid	Elev-gr	Lk gr	Lk nr	Ämne vecka 9, 10	Ämne vecka 11, 12	Elev-gr	Lk gr	Lk nr	Ämne vecka 9, 10	Ämne vecka 11, 12
8–9	1	1+3	1	I	5	8	2+4	28	VI	2
9–10	4	2+4	26	VIII	4	9	1+3	5	I	5
10–11	5	1+3	3	III	7	12	2+4	30	VIII	4
12–13	2	2+4	25	V	1	7	1+3	4	II	6
13–14	3	1+3	2	IV	8	10	2+4	29	V	1
14–15	6	2+4	27	VII	3	11	1+3	6	IV	8
8–9	6	2+4	39	V	1	11	1+3	18	II	6
9–10	3	1+3	14	II	6	10	2+4	41	VII	3
10–11	2	2+4	37	VII	3	7	1+3	16	IV	8
12–13	5	1+3	15	I	5	12	2+4	42	VI	2
13–14	4	2+4	38	VI	2	9	1+3	17	III	7
14–15	1	1+3	13	III	7	8	2+4	40	VIII	4

	Onsdag					Torsdag				
Tid	Elev-gr	Lk gr	Lk nr	Ämne vecka 9, 10	Ämne vecka 11, 12	Elev-gr	Lk gr	Lk nr	Ämne vecka 9, 10	Ämne vecka 11, 12
8–9	1	1+3	7	II	6	8	2+4	34	VII	3
9–10	4	2+4	32	V	1	9	1+3	11	II	6
10–11	5	1+3	9	IV	8	12	2+4	36	V	1
12–13	2	2+4	31	VI	2	7	1+3	10	III	7
13–14	3	1+3	8	I	5	10	2+4	36	VI	2
14–15	6	2+4	33	VIII	4	11	1+3	12	I	5
8–9	6	2+4	45	VI	2	11	1+3	24	III	7
9–10	3	1+3	20	III	7	10	2+4	7	VIII	4
10–11	2	2+4	43	VIII	4	7	1+3	22	I	5
12–13	5	1+3	21	II	6	12	2+4	48	VII	3
13–14	4	2+4	44	VII	3	9	1+3	23	IV	8
14–15	4	1+3	19	IV	8	8	2+4	46	V	1

Romerska siffror anger lektionsämne för mikrolektion 1
Arabiska siffror anger motsvarande ämne för mikrolektion 2
För siffrornas innebörd, se s 56

Undervisningstiden på 15 min tycks med hänsyn till experimentets målsättning vara mycket användbar. Den korta lektionstiden medför en ökad koncentration av undervisningserfarenhet och ger i förhållande till en vanlig 40 min lektion större möjlighet till ett exakt ihågkommande av händelserna i undervisningsprocessen. Den korta lektionstiden tvingar lkk att hålla sig till ämnet. Möjligheten till många olika undervisningsaktiviteter blir på ett na-

turligt sätt begränsad. Detta medför i sin tur också att handledningen kan koncentreras och blir lättare och mera enhetlig (jfr punkt 5.1.1).

Även för självkonfrontationsprocessen är den korta undervisningstiden av betydelse. Undersökningar har visat att längden för individens optimala uppmärksamhet eller koncentration vid uppspelning av videobandade lektioner sträcker sig ungefär till 25 min (Müller, 1966, s 39). Eleverna kan tänkas fungera olika beroende på den tidpunkt under dagen de blir undervisade. För att balansera sådana eventuella felkällor började inspelningen av t ex den första mikrolektionen första dagen med en lk från påverkningsgrupp 1 och 3 och nästa dag med en lk från påverkningsgrupp 2 och 4 etc (se tab 8). Har en bestämd lk däremot börjat sin första lektion t ex kl 8.00 så börjar lk också sin andra lektion kl 8.00 för att inte förändra betingelserna mellan första och andra mikrolektionen. I tabell 8 ges en sammanfattning av försöksplanen avseende den första mikrolektionen. Andra mikrolektionen upprepar exakt samma plan.

5.2 Beskrivning av designens olika nivåer

För en mera detaljerad utformning av den variansanalytiska designen (fig 2) kan man tänka sig flera alternativa modeller. Den detaljerade utformningen av modellen styrs framför allt av kombinationen mellan faktorerna T och H. Denna design förutsätter att faktorerna T och H påverkar lkk samtidigt. Men en *samtidig* påverkan kan utformas enligt åtminstone två alternativa grundmodeller:

Modell 1 skulle kunna vara att låta handledaren tillsammans med lkk se uppspelningen. Under denna tid ger såväl handledare som lkk muntliga eller skriftliga kommentarer till mikrolektionen. Denna modell skulle kunna kallas "*flätnings*"- eller "*saxningsmodell*".

Modell 2 skulle kunna vara att lägga båda påverkningsformerna blockvis efter varandra. Modellen skulle kunna kallas "*sekvensmodell*".

Utformningen av dessa två grundmodeller skulle i sin tur kunna leda till ett stort antal alternativa lösningar. Men för att diskussionen inte skall bli alltför detaljerad, koncentreras argumentationen endast på några få punkter. Diskussionen bygger på följande förutsättningar:

För att kunna bedöma sitt uppträdande och agerande i undervisningssammanhang skulle varje lk få till uppgift att hålla en mikrolektion på 15 min inför en elevgrupp på ca 10—15 elever. Hela denna lektion skulle videobandas i syfte att återges i sin helhet. Mikrolektionen skulle, enligt den traditionella handledningen (faktor H) sedan följas upp av en handledare (metodiklektor). Handledaren skulle vara närvarande under lektionen, eftersom så sker i den traditionella handledningen. Alla lkk skulle handledas av endast *en*

handledare för att undvika andra än intraindividuella differenser i fråga om handledningens utformning.

Med hänsyn till den externa självkonfrontationen via ITV/VR (faktor T) bestämdes att varje lk skulle hålla en mikrolektion på 15 min inför en elevgrupp på ca 10—15 elever. Hela lektionen skulle videobandas i syfte att återges i sin helhet. Under själva mikrolektionen skulle någon handledare ej vara närvarande. Handledaren skulle på så sätt ej ha något direkt eller indirekt inflytande på eleverna eller lkk. Efter lektionen skulle denna försöksgrupp sedan få tillfälle att se denna lektion tre gånger — första gången i omedelbar anslutning till lektionen, andra gången en dag senare och tredje gången sju dagar senare. Valet av antal uppspelningar och tidsintervall baseras på resultat från förförsöket (våren 1968). Alla lkk önskade vid detta tillfälle se sig själva flera gånger. Antalet gånger låg mellan 3 och 9 uppspelningar. De sex uppspelningstillfällena fram till tidsintervall h (se fig 3, s 63) ger en viss möjlighet till en mera systematisk granskning av hur stort antal uppspelningar som lkk bör få se vid ett studium av det egna undervisningsbeteendet. I den tillgängliga forskningslitteraturen rapporteras enbart att flera uppspelningstillfällen har förekommit, men antalet anges ej och inte heller vilken effekt detta har haft.

Diskussionen om en detaljutformning av den experimentella uppläggningen enligt modell 1 eller 2 bör föras med hänsyn till dessa förutsättningar. För att garantera att experimentets tolkningsmöjligheter ej går förlorade måste man uppmärksamma (1) *tidseffekt*, (2) *handledarens funktion* och (3) *sekvenseffekt*. För modell 1 och 2 kan anföras följande synpunkter:

Flätningsmodell (1)

Tidsfaktor:

Både handledare och ITV/VR påverkar lkk simultant under uppspelningstiden (15 min) dvs tidsperioden hålles konstant.

I vilken utsträckning lk kan utnyttja information från både handledare och ITV/VR, beror på interferenseffekternas storlek.

Handledarens funktion:

1. Hl får möjlighet att se ml två gånger.
2. De felkällor som uppstår vid traditionell handledning minskas eller förändras.
3. Hl måste följa ITV/VR-sekvensen och har därför ej tillfälle till mera grundläggande resonemang. Hl kan enbart ge några allmänna anmärkningar.
4. Hl kan distrahera lk i strävan att följa uppspelningen.
5. Hl och lk blir fixerade på detaljer.
6. Oberoende av om lk störs eller påverkas kan det ej urskiljas om förändringarna i lk perception och värdering beror på ITV/VR eller på handledningen, dvs allvarliga kopplingseffekter uppstår. Oberoende av vilka resultat som uppnås (positiva eller negativa) kan dessa ej tolkas.
7. Hl påverkan bör vara densamma för hela designen. Eftersom han ej är närvarande under T-faktorn, bör han ej heller vara närvarande under påverkningskombinationen (H+T).

Sekvensmodell (2)

Tidsfaktor:

Information som lkk tillförs via ITV/VR och den information som tillförs lkk via hl separeras, eftersom de inte kan anses vara identiska.

Detta medför att ITV/VR och hl skiljs från varandra. Men för att ge båda informationskällorna samma påverkningsmöjlighet måste påverkningstiden för denna modell ökas till 30 min.

Handledarens funktion:

1. Beroende på alternativa utformningar gäller punkt 1 och 2 för flätningsmodellen även för denna modell.
2. Hl bör ej separat få tillfälle att se ml via ITV/VR för att bevara tolkningsmöjligheterna.

Sekvenseffekt:

1. Att skilja hl från lk under uppspelningen medför sekvenseffekter.
2. Handledning bör följa omedelbart efter ml respektive efter den första videobanduppspelningen. Något större tidsintervall mellan båda skulle medföra att hl påverkas åtminstone av de övriga lkk prestationer. Dessutom motsvarar en sådan situation ej längre den traditionella handledningsmodellen.

Mot bakgrund av dessa argument valdes sekvensmodellen, eftersom enda fördelen med flätningsmodellen skulle vara att påverkningstiden kan kontrolleras. Tiden för handledarens och ITV/VR påverkan bestämdes till 30 min trots risken att det kan uppstå vissa tolkningssvårigheter på grund av den längre påverkningstiden jämfört med de övriga betingelserna.

Genom sekvensmodellen undviker man allvarliga kopplingseffekter och garanterar designens symmetri, dvs påverkan blir densamma för hela designen. Det kan naturligtvis diskuteras om det kanske vore bättre att låta handledaren

och ITV/VR simultant påverka lkk. Att en sådan koppling skulle kunna medföra viktiga pedagogiska och praktiska konsekvenser kan man inte utesluta eller bortse ifrån. Men då måste man på nytt ställa frågan om experimentets målsättning. — Är syftet att undersöka effekten av den traditionella handledningen i förhållande till effekten av självkonfrontation via ITV/VR? Vill man däremot undersöka vilka effekter som uppstår på grund av att ITV/VR och handledare är kopplade, så testar man hypotesen att denna kopplingseffekt medför en effektivisering av handledningen.

Skulle man vilja *integrera* även denna problemställning, uppstår ånyo problemet om sekvenseffekten. Lkk borde dessutom indelas i grupper med sex individer i varje grupp. Den sistnämnda hypotesen är dock en ny problemställning och kan ej lösas inom designens ram, därför att F-testen blir för svaga.

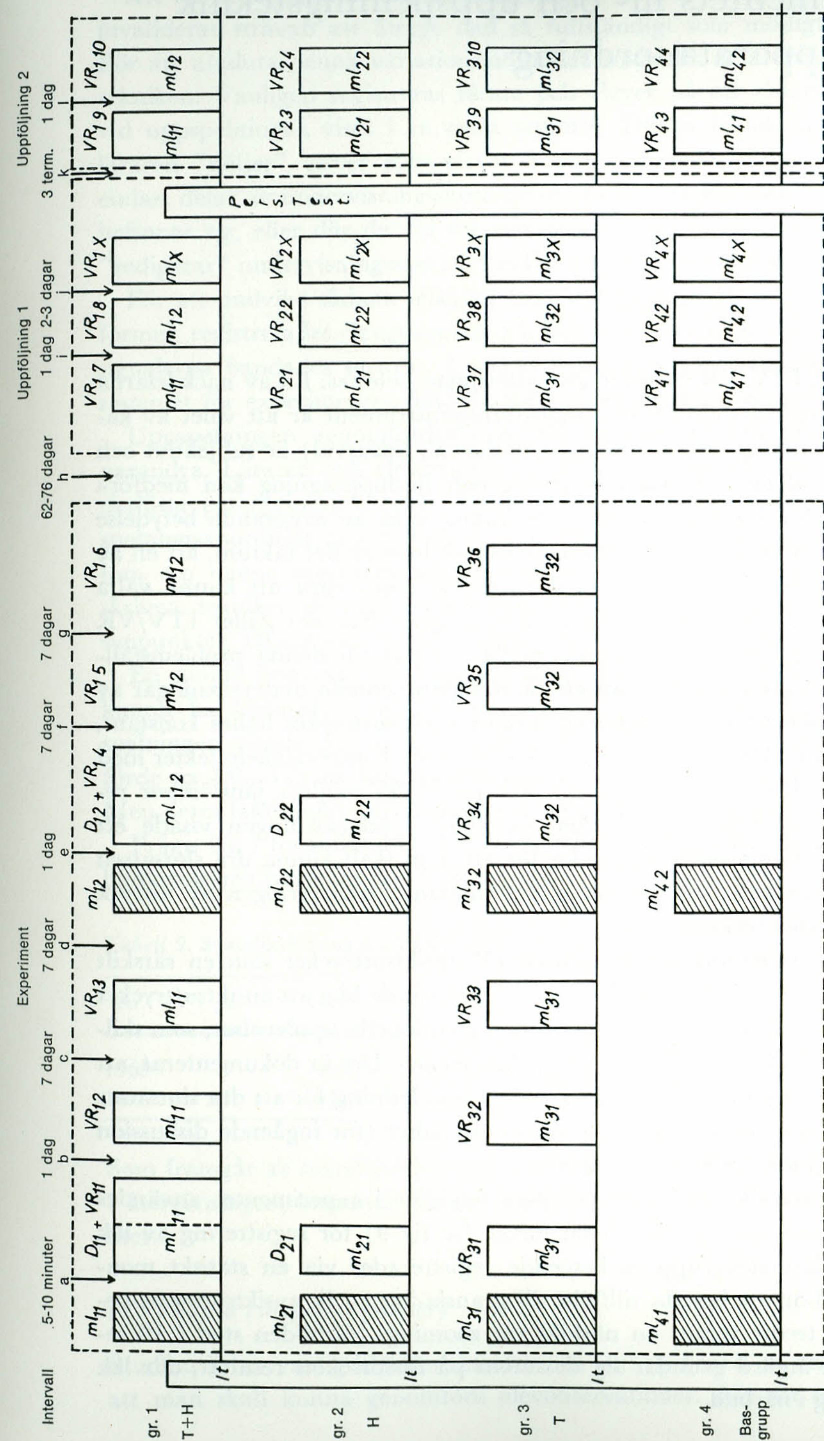
Att separera handledaren från lkk under ITV/VR-uppspelningarna medför sekvenseffekter med hänsyn till i vilken ordning ITV/VR- och handledarpåverkan kommer. Vill man kunna kontrollera sekvenseffekterna bör man i detta fall använda sig av ett slumpförfarande.

Slutligen skall nämnas att lkk i fält 4, dvs ingen påverkan alls, testades omedelbart efter att mikrolektionen avslutades, dvs tidsintervallerna mellan påverkningsformerna och testen har hållits konstanta för hela designen.

För att vi skulle kunna studera om en veckas undervisningspraktik inom malmöregionens övnings- och praktikskolor (som ingår i lärarutbildningen för andra terminens lkk) har någon effekt, fick lkk vid andra terminens slut se första och andra undervisningstillfällena under två på varandra följande dagar. Som sammanfattning av denna diskussion presenteras experimentets detaljerade utformning i figur 3.

VR_x innebär en uppspelning av en för fpp okänd lk. Denna uppspelning genomfördes för övrigt för samtliga lkk i terminskurs 2 för att möjliggöra ett studium huruvida det finns skillnader mellan olika grupper vid bedömningen av ifrågavarande lk. I anslutning till denna uppspelning administrerades ett antal personlighetstest. Under lkk sjätte termin vid LHM genomfördes en ny uppspelning av andra terminens inspelade mikrolektioner för att studera om lärarhögskoletiden har lett till förändringar i lkk perceptions- och värderingsstruktur.

En granskning av i vilken utsträckning (1) en veckas undervisningspraktik har haft någon effekt, (2) fpp vid bedömningen av VR_x skiljer sig från övriga lkk i terminskurs 2, (3) samband kan påvisas mellan experimentets resultat och sådana variabler, som är närmare specificerade genom testbatteriet (se kap 8) och (4) lärarhögskoletiden har lett till förändringar i lkk perceptions- och värderingsstruktur planeras att redovisas vid ett senare tillfälle.



Figur 3. Experimentets design

Förkortningar:

D11–12: Dyadisk konfrontation i samband med traditionell handledning, experimentgrupp 1 (första indexsiffran)

D21–22: Dyadisk konfrontation i samband med traditionell handledning, experimentgrupp 2 (första indexsiffran)

gr: Grupp

H: Experimentell faktor som anger påverkan traditionell handledning i form av dyadisk konfrontation.

lt: Inträdestest (grupptest)

ml: Mikrolektion. Första indexsiffran anger experimentgrupp, andra indexsiffran anger mikrolektion 1 respektive 2

ml_x: Mikrolektion X är en mikrolektion som har utförts av fp II under försöket vt 1968

Pers.test: Personlighetstest: testbatteri vid slutet av andra terminen

T: Experimentell faktor som anger påverkan externt förmedlad självkonfrontation via ITV/VR-tekniken

VR: Videobanduppspelningar. Första indexsiffran anger experimentgrupp, andra indexsiffran anger uppspelningstillfälle

6. Experimentets in- och uppspelningsteknik samt apparatanordning

Fördelarna med ITV/VR behöver inte ytterligare betonas. En av nackdelarna med ITV och videobandning som registreringsinstrument är att valet av kameravinkel eller redigeringsmetoder (spatialurval, tidsurval) är subjektivt och beroende av producenten. Även belysning och ljudupptagning kan medföra vissa problem. Att kameravinkeln skulle kunna vara av avgörande betydelse för skattningen av lk beteenden, framgår kanske bäst av det faktum, att en av porträttfotografens väsentligaste yrkesfärdigheter anses vara att kunna välja *rätt* vinkel för att fånga in individens *personlighet*. När det gäller ITV/VR föreligger såvitt bekant inga experimentella resultat till denna problemställning. Men det finns å andra sidan ett fåtal experimentella undersökningar av olika kameravinklar vid fotografering medan ansiktsuttrycket hålles konstant. Warr & Knapper (1968, s 307) undersökte ljus- och kameravinkeleffekter med hänsyn till stimuluspersonens framsida och profil. 96 möjliga jämförelser resulterade i 19 signifikanta skillnader ($\alpha = .05$). Undersökningen visade ett tillräckligt antal signifikanta skillnader för att man skall kunna dra slutsatsen att kameravinkeln kan vara en viktig determinant i genereringen av intryck genom fysionomiska tecken.

Det finns all anledning att ta hänsyn till ansiktsuttrycket som en särskilt viktig leverantör av information. Ekman (1950) visade bl a att ansiktsuttrycket förmedlar information om stimuluspersonens emotionella upplevelser, som skiljer sig från sådana förmedlade via kroppsliga tecken. Det är dokumenterat, att vi kontinuerligt utnyttjar fysionomiska tecken som ledning för att dra slutsatser om andra personers tankar, känslor och personlighet (för ingående diskussion se Warr & Knapper, 1968, ss 296—299).

För att utesluta eller balansera möjliga felkällor i experimentet användes två kameror sammankopplade via en mixer (se fig 9) för registrering av lkk beteenden, medan elevgruppens beteende registrerades via en statiskt monterad kamera. För att bereda tillfälle till granskning av lk ansiktsuttryck, registrerades var tredje minut en närbild via zooming. Närbilden stod i 10 sekunder. Denna åtgärd grundar sig dessutom på förförsökets resultat, där lkk önskade få se sig i närbild.

Att en pedagogisk-metodisk redigering av undervisningsprocessen skulle ha invaliderat strävan att återge den så fullständigt som möjligt är uppenbart. För att utesluta denna variationsmöjlighet användes den beskrivna 3-kamera-tekniken. Vanligen registreras lärare och elever på ett sådant sätt, att båda vid uppspelningen visas i en enda monitor. Denna teknik har nackdelen att läraren "döljer" elever eller elever åtminstone delvis döljer läraren och att endast delar av undervisningsskeendet inspelas, antingen den delen där läraren befinner sig, eller där de vid ett visst tillfälle aktiva eleverna finns, dvs man "redigerar" undervisningsskeendet enligt mer eller mindre explicita kriterier.

För att undvika sådana felkällor som är förknippade med olika redigeringsformer, registrerades elevgruppens och lärarens beteenden synkront, men videosignalerna bandades separat. I blockschemat (fig 9, s 70) över det tekniska systemet för experimentet, åskådliggörs tillvägagångssättet grafiskt.

Uppspelningen genomfördes över två monitorer, som placerades bredvid varandra. Läraren och elevgruppen visades synkront. Denna teknik antas underlätta perceptionsprocessen. En experimentell undersökning av denna uppspelningsanordning ingår dock ej i undersökningen. Det förefaller emellertid som om denna registreringsteknik är mera tillfredsställande än tidigare beskrivna tekniker dels ur forskningsmetodiska synpunkter, dels ur individens synpunkter. Dessutom bortfaller experimentatorfaktorn som en möjlig felkälla.

En möjlig irritation vid uppspelningen skulle i samband med denna teknik kunna ha förorsakats av synkroniseringsproblemet. Vid den synkrona uppspelningen visade det sig nämligen, att en 15 min uppspelningsperiod medförde ca 1.0—1.5 sek fördröjning, dvs en av apparaterna går något snabbare. Men detta faktum syns ej ha påverkat lkk bedömning av undervisningsskeendet.

Fråga 2 a i formulär FC (jfr Bierschenk, 1972 e, gul bilaga 5) lydde: Har bildkvaliteten påverkat Din bedömning?

Tabell 9. Svarsfördelning för fråga 2a

År	Ja	Nej
1969	6	42
1970	7	41

Som framgår av svarsfördelningen har inte alla lkk varit helt ostörda av brister i återgivningen, beroende på att denna tyvärr blir av varierande teknisk kvalitet.

ITV-studio (inspelningslokal)

Studion är av vanlig klassrumsstorlek och utrustad med en one-way-screen för att man skall kunna genomföra elevobservationer. För att möjliggöra ITV-

inspelningar av undervisningssituationer och för att åstadkomma en användbar ljudupptagning kompletterades det ursprungliga klassrummet med mattor och gardiner. Vid genomförandet av experimentet hade ITV-studion den utrustning och det utseende som framgår av figurerna 4 och 5. Följande apparatur kom till användning:

Kameror

3 st Philips fjärrstyrda kameror med vidikonrör.
Kamera 1: EL 8010/20, Canon zoomobjektiv TV 16, 25–100 mm, 1:1.8, placerad 1.8 meter över golvet.
Kamera 2: EL 8010/20, Canon zoomobjektiv TV 16, 25–100 mm, 1:1.8, placerad 1.8 meter över golvet.
Kamera 3: EL 8010/20, Canon zoomobjektiv, TV 16, 25–100 mm, 1:1.8, placerad 2.2 meter över golvet.

Kamera 1 och 2 kopplades som ett 2-kamera-system med bildväxling över videomixer, medan kamera 3 installerades med fast position för registrering av elevgruppen.

Belysning

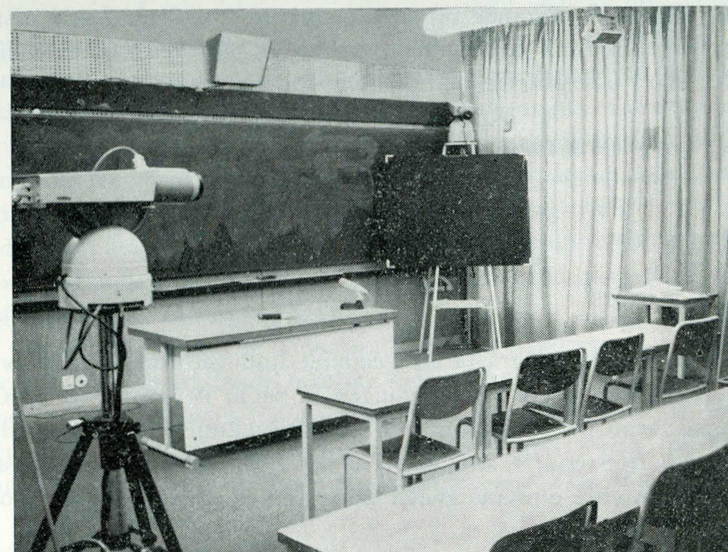
Utom den normala klassrumsbelysningen användes 4 st halogenstrålkastare à 1 kW riktade mot taket för att ge indirekta belysningseffekter. Klassrummets ljusstyrka uppmättes till 800—1 000 lux. Den normala belysningen i ett klassrum uppgår (Philips Ljusboken, 1966) till 300—500 lux.

Ljudupptagning

För inspelning av elevernas ljud användes 3 st takmikrofoner (för position se fig 4) och för inspelning av lärarens ljud användes en hals- och en bordsmikrofon.

Ljudupptagning

Mikrofon	*Typ*	
Fabrikat Sennheiser		
I	MD 421	Placerad 1.8 meter över golvet (närmast tavlan).
II	MD 421	Placerad 1.6 meter över golvet.
III	MD 421	Placerad 1.6 meter över golvet.
1	MD 1008	Är en trådlös UKW-sändare/mikrofon som fp har om halsen.
2	MD 421	Bordsmikrofon, placerad på katedern.



Figur 4. Apparatanordning i ITV-studion



Figur 5. Apparatanordning i ITV-studion

Ljudupptagning för lkk genomfördes 1968 dels med mikroport, dels med bordsmikrofon. Då mikroporten visade sig otillförlitlig (variationer i signalstyrka och störningar) användes år 1969 uteslutande bordsmikrofonen. Lkk fick dock även 1969 bära UKW-mikrofonen för att inte förändra den experimentella situationen. Att åstadkomma en tillfredsställande ljudkvalitet i en

klassrumsstudio är fortfarande ett mycket svårlöst problem. I vad mån ljudkvaliteten påverkade lkk bedömningar, anger lkk på fråga 2 b i formulär FC, som lydde: Har ljudkvaliteten påverkat Din bedömning?

Tabell 10. Svarsfördelning för fråga 2b

År	Ja	Nej
1969	5	43
1970	12	36

Medan år 1969 endast några få lkk bedömde ljudöverföringen vara av så låg kvalitet, att den har påverkat bedömningen, är det är 1970 avsevärt flera. Detta kan knappast bero på att vi kopplade in bordsmikrofonen, eftersom den gav ett påtagligt bättre resultat. Möjligen har lkk 1969 haft en större toleransnivå genom allmänna yttre omständigheter kring den experimentella situationen.

Kontrollrummet

Den apparatur som var nödvändig för upptagning av video- och ljudsignalerna samt för återgivningen placerades i kontrollrummet. Kontrollrummets utseende visas i figur 6.

För inspelningen användes följande videobandapparater: 1 Ampex VR 7003 och 1 Ampex VR 7000. Beträffande tekniska data se Ampex manuals (1969). För upptagningen av videosignaler användes videoband fabr Memorex och Scotch. För att kunna synkronisera båda videomaskinerna exakt spelades vid



Figur 6. Apparatanordning i ITV-kontrollrummet

början av varje 15 min lektion in dels en testbild, dels en 20 sek ljudsignal. Varje inspelning avslutades med en testbild. Dessutom interfolierades inspelningarna med ljudsignaler (1 min intervall) som ligger på lkk videoband. Både klassens och lkk ljud ligger däremot på elevgruppens videoband.

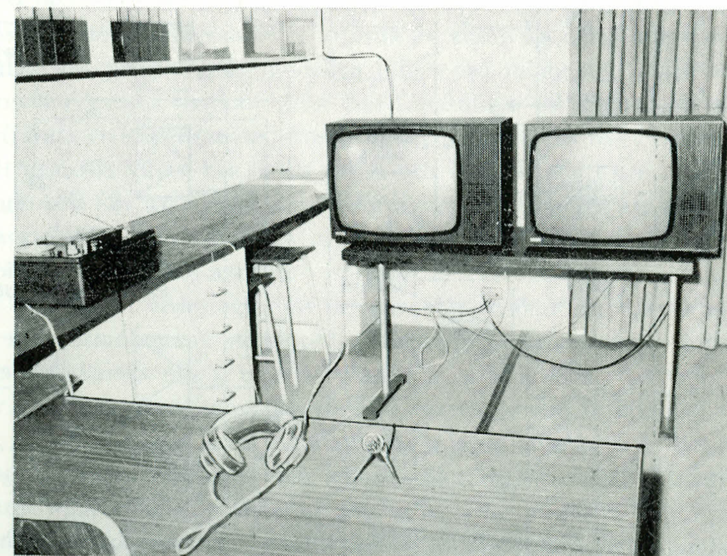
För att vi skulle kunna bearbeta det verbala materialet utan att i sådana sammanhang för varje gång behöva spela upp videobanden, har klassens ljud dessutom spelats in på separat ljudband. 4 st Tandberg typ 14 och 82, utrustade med 2 separata ingångar, kom till användning för upptagning av kommentarer och tidspulser samt originalmikrofon. För upptagning av ljudsignalen användes ljudband, fabr Agfa PE 41.

För uppspelningen av de videobandade mikrolektionerna användes 2 st Ampex VR 5103. Beträffande tekniska data se Ampex manual (1969). Ljudbandspelare som användes för upptagning av simultankommentarerna under självkonfrontationen, placerades under experimentets första fas i uppspelningsrummen, men flyttades under experimentets andra fas till kontrollrummet.

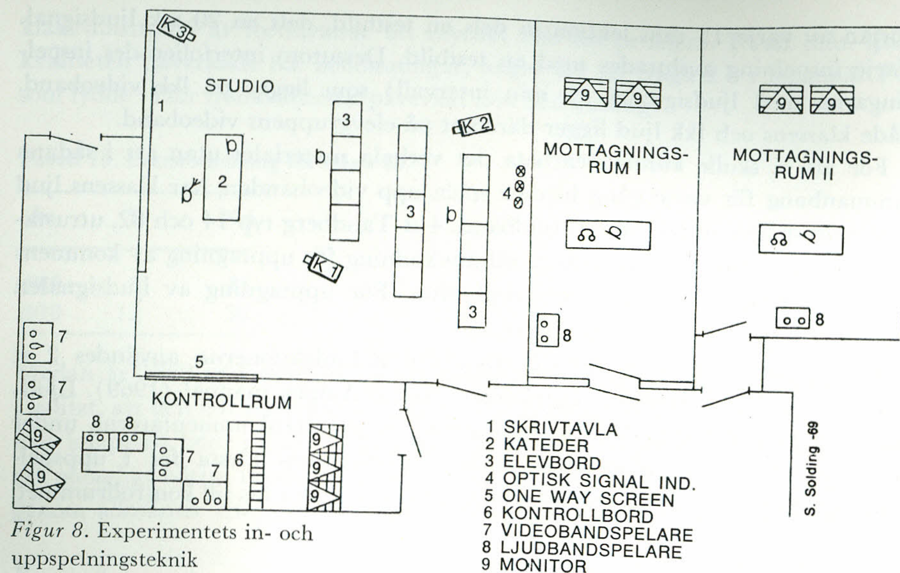
Mottagningsrum I och II (återgivningslokal)

För självbedömning och bedömning genom en bedömningspanel användes 6 st monitorer (AGA, M 2301, 23") som matades med videosignaler från VR 5103 bandmaskiner och utan separat distribution vid uppspelningen. Mottagningsrummens utseende och monitorernas placering visas i figur 7.

För ljudmottagning vid uppspelningen användes hörlurar av typ ELEGA DR 59, medan upptagningen av simultankommentarerna skedde via bords- respektive stativmikrofon.

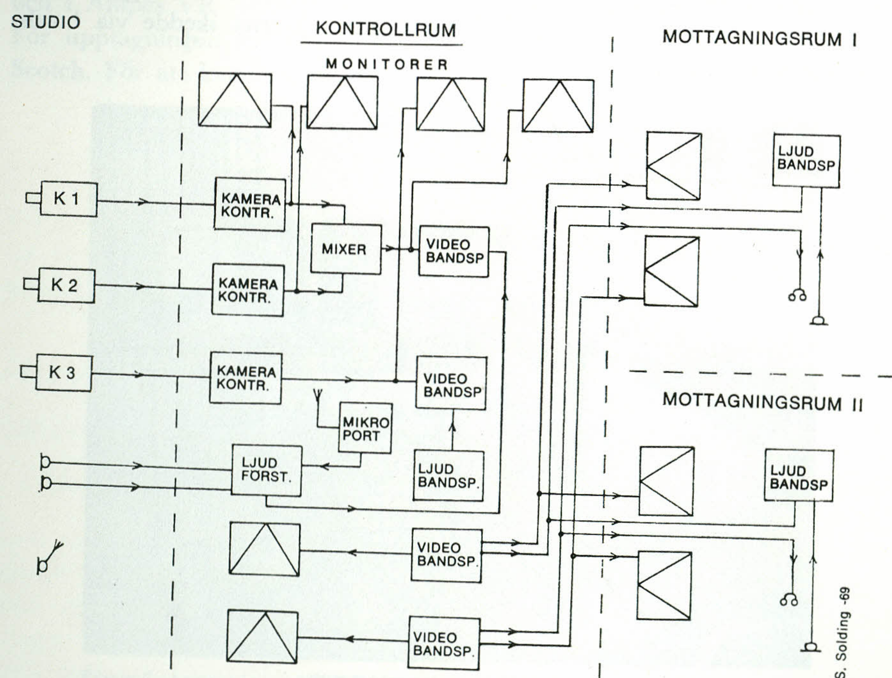


Figur 7. Apparatanordning i ITV-mottagningsrum



Figur 8. Experimentets in- och uppspelningsteknik

I figur 8 presenteras ett schema över lokaler och apparatanordning för videobandning och återgivning av mikrolektionerna. Den tekniska utrustningen för experimentet presenteras i form av ett blockschema i figur 9.



Figur 9. Blockschema över den tekniska ITV/VR-utrustningen för experimentet

7. Självbedömning och expertbedömning

7.1 Rollen som "extern observatör-kommentator" av sig själv

Genom att lk via ITV/VR kan åstadkomma en extern perception och värdering av sin egen undervisning, blir han sin egen "externa observatör och kommentator". Forskningen kring perception och analys av interaktionsprocesser i en undervisningssituation tillförs med denna observationsteknik ett helt nytt element, nämligen en objektiv återgivning av information om individens beteenden genom en "extern självdistansering i tid och rum". Den externa självdifferentieringen, dvs en distinkt avgränsning av individens själv från undervisningsprocessen, medför en ny form av självdiagnos och självvärdering, jämfört med t ex Cooleys (1968, ss 87—91) "looking-glass-me".

Vid studiet av interaktionsprocessen i en undervisningssituation har använts många olika metoder för att observera, registrera och kategorisera lärar- och elevbeteenden (jfr Medley & Mitzel, 1963, ss 267—328, Withall & Lewis, 1963, ss 683—714 och Bjerstedt, 1970 b). Men alla dessa observationssystem baseras på externa observationer utförda av andra än den undervisande själv. Med den här presenterade modellen införs nya moment i lärarutbildningen, som avviker ganska markant från den förefintliga traditionen, nämligen (1) att befrämja en experimentell attityd gentemot lärarens roll i undervisningen, dvs att utveckla viljan hos lk att pröva och utvärdera sitt eget undervisningsbeteende och (2) att bygga upp en färdighet och förmåga att använda sig av feedbacktekniker, dvs objektiva metoder som ITV/VR, för att analysera det egna beteendet och dess resultat.

Det är allmänt vedertaget, att externa observatörer kan jämfört med individen själv åstadkomma större objektivitet vid beskrivning av en persons beteende. Men kanske missar man därvid viktig information.

När handledaren tillsammans med lk bearbetar vad lk iakttagit eller själv utfört, är båda tvungna att i stor utsträckning analysera och diskutera utan tillgång till mera objektiva och systematiska registreringar. Handledarkommentarer upplevs därför ofta som subjektiva eller orättvisa impressioner. Utbildningen av lkk tycks vara baserad på ett av nästan alla lärarutbildare accepterat antagande: lärarens eller lk handlande är beroende av vad de tror är

viktigt och vad som är viktigt bestäms, som påpekades i kap 1, av individens referensram.

Men hur mycket vet vi egentligen, på mera objektiva grunder, om lk sätt att se sig själv och sitt eget beteende i en undervisningssituation?

7.2 Pedagogiska experters roll som bedömare av lärarkandidater

En av pedagogiklektorns huvudfunktioner i lärarutbildningen är att åstadkomma (åtminstone teoretiskt sett) en integration av teori och praktik. För att bedöma den studerandes utbildningsframgång och yrkesskicklighet, observerar vanligen en eller flera pedagogiska experter lk i en undervisningssituation. Bristen i denna utbildningsform är dock den stora klyftan mellan teori och praktik. Lk skjutsas igenom sin utbildning mellan en allmän teori och en relativt ostyrd praktik. Den i många sammanhang åberopade integrationen överlåtes däremot i stor utsträckning till lk själv (jfr Bjerstedt, 1970 b, ss 212—238). Eftersom forskningsresultaten angående lärarskattning och lärarlämplighet hittills ej har kunnat ge någon lämplig utgångspunkt för mera detaljerade rekommendationer (se Sundgren, 1967), borde utbildningen planeras så, att lk successivt kan pröva ut och träna in sådana beteenden, som är effektiva i hans eget fall, dvs utnyttja sin egen person som kontroll- och regleringsmekanism.

Att tänka sig att katalogisera lösningar för alla de undervisningsproblem, som läraren kan tänkas komma att möta, är en metod, som faller på sin egen orimlighet. Vad som saknas för att man skall kunna utbilda en dynamisk lärarpersonlighet är uppenbarligen länkar mellan teoretiska beteendevetenskapliga kunskaper och själva undervisningssituationen. Att på det nuvarande sättet studera lärarens undervisningsskicklighet och specifika undervisningsmetoder har ej visat sig vara särskilt framgångsrikt. Ett allt större antal pedagoger poängterar, att lärarens personlighet, självuppfattning och attityder måste betraktas som "modifierare" av lärarens effektivitet. Öppenhet, förmåga att reagera på ett adekvat sätt i en situation och värdera situationen med hjälp av realistiska (objektiva) fakta i situationen, har visat sig vara direkt relaterade till framgången i undervisningen (Combs, 1965, s 78).

Bills, Macagnoni & Elliot (1964, s 1) sammanfattar ett antal forskningsresultat på detta område. De skriver:

"there appears to be direct relationship between the openness of a teacher to his experience, both past and present, his judged teaching success, his effect on attitudes toward self and others, the locus of responsibility for decision making within his classroom, his ability to change in a learning situation, and the quality of the helping relationship he offers pupils."

Intern television, videobandning och mikrolektioner är nya undersökningstekniker, som möjliggör ett nytt angreppssätt för att undersöka på vilket sätt man skulle kunna åstadkomma en bättre integration av lärarutbildningens teoretiska och praktiska aspekter.

För lärarutbildningen bör det vara av vitalt intresse att kartlägga på vilket sätt lärarutbildaren percipierar och värderar lk i en undervisningssituation jämfört med den studerandes egna perceptioner och värderingar för att kunna planera en målinriktad utbildning. För att lärarutbildaren skall kunna påverka lk i en eller annan riktning, måste denna tydligen låta sig "förändra" i önskad riktning. Detta kräver en höggradig individualisering av träningsprogrammen och en ingående översyn av pedagogik- och metodiklektorns funktioner i lärarutbildningen. Handledning är dessutom ett tidskrävande arbete, och de flesta handledare lider av ständig tidsbrist. På detta sätt riskerar man att ett mycket ömtåligt område i lärarutbildningen försummas. Få lkk har därför chansen att få en adekvat träning för att utveckla sitt undervisningsbeteende inom ramen för respektive individs personliga karakteristika.

Cohen, L. (1969, s 48) utförde en undersökning rörande *functional dependence, exchanges, and power of influence* i syfte att kartlägga lk situation under övningsundervisningen eller praktikperioden, med särskild tonvikt på handledarens roll. Beträffande pedagogik- och metodiklektorer konstaterar han, att deras korta och oregelbundna besök hos lkk gör lektorerna "organizationally distant, infrequent in communication, and it will be argued, less constrained to attempt to influence the student than the class teacher".

Cohens undersökning visar, att läraren på respektive skola har en större påverkan på lk utformning av beteende och attityder under praktikperioden än lektorerna från lärarhögskolan. De relationer som existerar mellan lk och de olika lärarutbildarna tyder på att lärarhögskolan visserligen fortfarande är det ställe som påverkar lk föreställningar om ett ändamålsenligt lärarbeteende, men praktikskolan representerad av klassläraren "is in a powerful position to effect change in those conceptions during the relatively short time that students are exposed to the practical problems of the classroom" (Cohen, L., s 50).

Det investeras betydande resurser i form av finansiella och personella tilldelningar i lärarutbildningen för att förbättra undervisningen. Men vad vet vi egentligen om lärarutbildningens effekt på den studerande? Trots alla motsägelsefulla forskningsresultat har expertskattningarna som utbildnings- och värderingsinstrument alltjämt en stark ställning inom lärarutbildningen. För att olika påverkningsteknikers effekter skall kunna utvecklas och kontrolleras, borde det vara av centralt intresse, att expertskattningarna blir föremål för experimentella studier.

Experimentets mikrolektioner bedömdes av fyra pedagogiklektorer. För urvalet av personer till bedömningspanelen skulle följande kriterier vara uppfyllda: (1) bedömaren har själv genomgått lärarutbildning och förvärvat

kompetens för att undervisa på mellanstadiet, (2) bedömaren har flera års undervisningserfarenhet på detta stadium, (3) bedömaren har en avslutad beteendevetenskaplig skolning och (4) bedömaren har egen beteendevetenskaplig forskningserfarenhet. Dessa förutsättningar borde kunna ge bedömaren möjlighet att skatta lk undervisningsprestationer enligt pedagogisk-psykologiska och praktiska principer. För experimentet skulle bedömningspanelens skattningar kunna betraktas som jämförelsekriterium, dvs som "norm" för att t ex kunna definiera i vilken utsträckning fpp självskattningar är relaterade till gällande pedagogiska principer, dvs till expertpanelens skattningar.

En annan frågeställning skulle kunna vara att undersöka på vilket sätt självskattningar och expertskattningar kan kombineras. Att låta lkk retrospektivt skatta den egna undervisningsprestationen med något konventionellt observationsinstrument medför att lkk utgångsläge knappast kan jämföras med t ex pedagogik- eller metodiklektorns. Medan lk är tvungen att koncentrera sin uppmärksamhet i första hand på sina uppgifter som undervisare kan lektorerna helt ägna sig åt observationsuppgiften. De observationsdata som framkommer kan (åtminstone av observationsmetodiska skäl) knappast anses vara jämförbara.

Undervisnings-situation	Tids-intervall	Uppspelning via ITV/VR	Perception
XOY XOX	Flexibel	XOY XOX distansering ⟷	XO ⟨XOY XOX⟩
XOY XOX	Flexibel	XOY XOX distansering ⟷	ZO ⟨XOY XOX⟩

Symboler

O: operation
X: organism, vilken som helst
Y: objekt (organism eller materia)
Z: extern observatör

Figur 10. Undervisning och självdistansering via externt förmedlad självkonfrontation

Symbolerna i uttrycket XOY anger en subjekt-objekt-relation. Den utläses från vänster till höger så att första ledet "X" är den agerande mot sista ledet "Y", som är objekt för en operation. Om vi byter ut Y mot X, så innebär uttrycket XOX att den agerande organismen X utför en operation mot sig själv. Med detta paradigm antas undervisningsprocessen, första delen i figur 10, fullständigt definierad (för utförligare beskrivning se Bierschenk, 1972 f, ss 23–33).

Genom ITV/VR-tekniken kan lk åstadkomma en extern perception och värdering av den egna undervisningen och således bli sin egen "externa observatör och kommentator". Denna roll motsvarar onekligen mycket mera pedagogik- eller metodiklektorernas roll. Både lk och lektorerna får på så sätt samma mängd av audiovisuell information för sin bedömning. Dessutom borde lkk och pedagogisk-metodiska experters bedömning av sådana videobandade lektioner underlätta en etablering av en gemensam referensram och användning av gemensamma bedömningskategorier.

Rollen som observatör och kommentator medför en distinkt avgränsning i tid och rum från själva undervisningsprocessen. Lk är alltså en del av den ursprungliga interaktionsprocessen, inom vilken lk utför många olika operationer, som sedan i lk roll som extern observatör och kommentator blir föremål för perception och värdering, förmedlad via ITV/VR. Sammanfattningsvis presenteras i figur 10 i grafisk form lk roll som extern observatör-kommentator i relation till bedömningspanelen.

7.3 Bedömningsförfarandet

För självbedömningen och panelbedömningen användes skattningsskalor. I användningen av skattningsskalor utförs bedömningen som ett slags betygssättning i förhållande till vissa icke omedelbart givna normer. Vid användningen av sådana skalor måste bedömaren ha kännedom om ett antal individer, som skall bedömas i det särskilda fallet och mellan vilka jämförelser skall ske. Likheten mellan experimentets mikrolektioner och lk kommande praktik, dvs den undervisningssituation som pedagogik- och metodiklektorer vanligen bedömer, framgår av studieplanernas anvisningar för terminskurs två (M 2). Enligt dessa skall lk i M 2 endast undervisa 20—30 min av en lektion. Därefter övertar handledaren eller någon annan lk undervisningen. För att vi skulle kunna fixera bedömarnas bedömningsnormer så långt som möjligt tränades panelen enligt ett skattnings- och värderingsschema (F III) i videobandade undervisningssituationer, som hade inspelats under förförsöket. Bedömningar, baserade på skattningsskalor, medför välkända risker och felkällor (för utförlig diskussion, se t ex Ekman, 1950).

I en jämförelse av effektivitet och noggrannhet i skattningar av visuella, auditiva och skriftliga protokoll för bedömning av lärarens undervisningseffektivitet (avseende att förklara och interpretera) fann Unruh (1967, s 115) audio-visuell återgivning överlägsen alla andra protokoll och kombinationer av protokoll ($\alpha = .01$) i exakthet i bedömning. Effektiviteten skattades med hänsyn till elevernas prestationer. Som förklaring till detta resultat anges att det audiovisuella protokollet möjligen bäst motsvarar själva klassrumssituationen och att bedömarna var mest förtrogna med bedömningen av lk effek-

Tabell 11. Panelens bedömning av mikrolektionerna (tidsschema för antal timmar och antal bedömningar)

Tidpunkt Månad/dag	Bedömning år 1969 tim	antal	Bedömning år 1970 tim	antal	Verksamhetsform	
03–20	2	2	4	2		
21	2	2				
22						
23						
24			4	2	Bedömarträning	
25	2	2				
26						
27	2	2				
28	2	2				
29						
30					Principbedömning	
31	8	14	4	7		
04–01	8	14	3	6		
02	8	14	4	8		
03			3	6		
04						
05						
06						
07			4	8		
08	Påsk		3	6		Bedömning av experimentets lektionsmaterial i dess helhet
09			7	13	Principbedömning (70)	
10			7	14		
11						
12						
13			3	7		
14	6	12	4	7	Principbedömning	
15	8	16				
16	4	7	4	8		
17	7	14	4	7		
18	7	14	4	8	Rebedömning 1 (n = 9)	
19					Principbedömning	
20						
21						
22	7	14				
•						
•						
•					Principbedömning (70)	
06–09	7	15	7	15	Rebedömning 2 (n = 30)	
10	7	15	7	15	Principbedömning (70)	

tivitet i klassrumssituationer. Enligt denna undersökning borde man använda sig av videobandade lektioner vid bedömningen av lk undervisningprestationer. Warr & Knappers (1968, ss 302—308) resultat pekar i samma riktning. Ett tillägg av fotografisk information till en verbal beskrivning förändrar påvisbart ($\alpha = .05$) bedömarens perception av stimuluspersonen.

För att fixera bedömningspanelens referensram mot bakgrund av skattnings- och värderingsinstrumentet och för att på så sätt förbättra skattningsskalornas diskrimineringsegenskaper samt instrumentens reliabilitet genomfördes träningsbedömningar med anslutande diskussionspass. Hela förfarandet vid bedömning av experimentets mikrolektioner redovisas i tabell 11. Bedömarna som ingick i panelen var samma personer både 1969 och 1970. Som tabell 11 visar, skiljer sig fördelningen av bedömningarna för år 1969 och 1970. Denna variation i såväl tidslängd per dag som över hela bedömningsperioden (20/3—9/6) är betingad framför allt av schematekniska skäl.

Som framgår ur tabell 11 omfattade bedömarträningen 8 tim. Under träningsperioden utfördes kontinuerliga kontroller av bedömarnas överensstämmelse. Efter träningsbedömningar utförde varje bedömare en "principbedömning". Med denna beteckning avses en skattning utan uppspelning av ett videoband. Bedömarna uppmanades att anta för item (a) i formulär F III skalans mittposition och sedan bedöma värdet i item (b). Syftet med denna skattning är att få ett värde på bedömarens mera generella värdering av de i F III ingående beteendeaspekterna i relation till experimentets mikrolektioner. För att kunna studera utvecklingen i bedömarens mera generella bedömning utfördes tre principbedömningar, en i början, en när ungefär hälften av materialet hade bedömts, och en i anslutning till rebedömning 1.

Ordningsföljden för uppspelning av de videobandade lektionerna (1 och 2) bestämdes via en slumptabell. Av de fyra pedagogiklektorer som ingick i bedömningspanelen bedömde nr 1 och 2 hela materialet, medan nr 3 bedömde ena hälften och nr 4 andra hälften av materialet. Samtliga utförde sina bedömningar i relativ isolering.

Rebedömning 1 avser 9 lektioner som slumpades ut ur de första 20 bedömningarna. Efter sex veckor (9/6—10/6) genomfördes rebedömning 2, avseende 30 mikrolektioner som utgör ett slumpmässigt stickprov ur hela lektionsmaterialet. Panelen utförde sina bedömningar under samma betingelser som lkk.

8. Mätinstrument och mätningstillfällen

Vid användningen av ITV/VR-tekniken som utbildnings- och forskningsinstrument faller man i det stora flertalet av undersökningar tillbaka på känd testmetodik. Men vill man utreda den ursprungliga problemställningen, dvs undersöka effekten av ett nytt medium, kan man för det mesta ej använda gamla test på den nya frågeställningen. Vill man t ex undersöka nya personlighetspsykologiska problem med den gamla arsenalen av test, då har man övergett sina första intentioner och söker svar på frågor som testkonstruktören en gång ställde. Även inom den pedagogisk-psykologiska forskningen måste vi ha rätt att låta forskningsprocessen fortskrida genom en serie stadier, där de första ofta är av explorativ natur, dvs trevande och diffusa. I detta första stadium spelar fria observationer, autoexperiment, osystematiska intervjuer och samtal en stor roll. Tyvärr har man i den hittills genomförda forskningsverksamheten kring ITV/VR-tekniken i allmänhet förbigått denna fas.

Vid utvecklingen av mätinstrument för experimentet började vi från "scratch". Konstruktion och urval av mätinstrument baserades på följande frågeställning: "Vad är det egentligen som vi önskar att våra fpp skall tala om för oss?" Det är allmänt vedertaget, att extrospektion kan leda till större objektivitet vid beskrivning av en persons beteende. Att mätningar som bygger på individens introspektion och mätningar som bygger på panelbedömningar ofta skiljer sig avsevärt från varandra, beror kanske i ganska stor utsträckning enbart på skillnaden mellan de variabler som kommer till användning för att beskriva ett beteende. Kanske är skillnaden i det objektiva medvetandet om de beteendeaspekter, som blir explicita, inte alls så stor. Här skulle ITV/VR-tekniken kunna användas för konstruktion av sådana mätinstrument som innehåller variabler som både individen själv och bedömningspanelen använder för beskrivningen av ett visst beteende.

Utvecklingen av mätinstrumenten för registrering av fpp reaktioner på självkonfrontationsupplevelser påbörjades med det tidigare nämnda förförsöket (vt 1968). Sex lkk från lågstadielärarutbildningslinjen genomförde ett antal mikrolektioner med varierande tidslängd (10, 15, 20 min). Sedan fick dessa fpp se sina lektioner via ITV/VR. I samband med uppspelningen kommenterade fpp sina lektioner simultant. Dessa kommentarer spelades in på ljudband. Genom en impressionistisk innehållsanalys utvecklades ett antal kategorier

som bildade utgångsmaterialet för konstruktion av ett antal items. Dessutom genomfördes i anslutning till förförsöket individuella intervjuer samt gruppintervjuer som gav ytterligare information om vilka items som skulle kunna ingå i mätinstrumenten. (För diskussion av de principer som styrde kategoriseringsarbetet, se Bierschenk, 1972 f.)

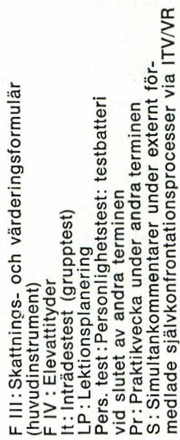
I de flesta undersökningarna används sjugradiga skattningsskalor. Men även skattningsskalor med annat antal (5, 9, 12, 11, . . ., 20) skalsteg används. Dessutom kan man naturligtvis också använda grafiska skalor, dvs ersätta talföljden med en vågrät linje. Bedömaren kan markera vilken punkt som helst på linjen och är sålunda inte bunden till ett begränsat antal grader. Warr & Knapper (1968, ss 109—114) utförde ett antal experiment för att undersöka reliabiliteten av olika skallängder i samband med semantiska differentialer. Resultatet blev att skalor med nio eller flera skalsteg blev lika tillförlitliga som en sjugradig skala. En transformering av t ex niogradiga skalor till sjugradiga medförde dessutom nästan identiska profiler. Även grafiska skalor gav samma resultat. Fördelen med en sjugradig skala framför en femgradig skala är att den sjugradiga skalan tillåter en normalisering av påtagligt sneda fördelningar vilket i sin tur tillåter ett parametriskt testförfarande. För- och nackdelar med den grafiska skalan diskuterades också men den förkastades på grund av den mycket större arbetsbelastningen vid utvärderingen, som jämfört med en intervallindelad skala ej leder till mer eller bättre information. Vi bestämde oss alltså för att använda sjugradiga skattningsskalor för experimentet. Varje beteendeaspekt som ingår i mätinstrumenten skattas med hjälp av en sjugradig skattningsskala.

Beslutet om vilka items som skall ingå i ett mätinstrument, måste fattas på arbiträra grunder, även om förförsöket gav impulser till urvalet. Bakom den praktiska problemställningen finns det en mycket stor teoretisk osäkerhet, nämligen frågan: Hur kan vi veta att ett mätinstrument verkligen mäter perception och värdering när det ej finns någon exakt och allmänt accepterad definition av denna process?

Att definiera "hela perceptionsprocessen" är logiskt sett omöjligt och eftersom vi inte kan specificera perceptionsprocessen fullständigt, vet vi inte heller om vi har ett mätinstrument av full validitet.

Trots att vi inte kan veta om vi har mätt alla aspekter av den psykologiska processen, som vi är intresserade av, så kan vi dock ställa frågan: Vilka av de olika mätningarna är valida och reliabla? Att besvara denna fråga ingår i experimentets problemställning. En lämplig väg för att reducera osäkerheten i mätinstrument är användningen av olika dimensionsanalytiska tekniker.

En översikt över mätningarna och mätningstillfällena under experimentets olika faser ges i figur 11.



Figur 11. Mätning och mätningstillfällen (för staplarnas innebörd se fig. 3)

8.1 Modellen för konstruktion av mätinstrument

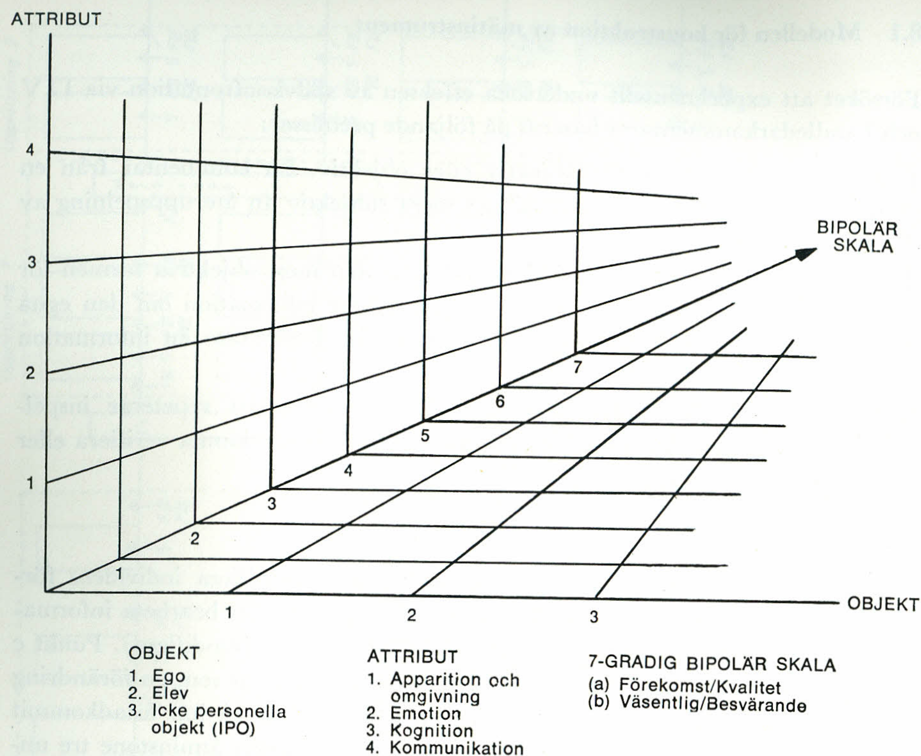
Försöket att experimentellt undersöka effekten av självkonfrontation via ITV och handledarkommentarer baseras på följande premisser:

1. Konfrontation kan vara subjektiv eller objektiv. En kommentar från en handledare, t ex "du ser arg ut", är mera subjektiv än återuppspelning av ett videoband.
2. Återgivning av videobandade beteenden är den mest objektiva formen för en objektiv konfrontation, eftersom den tillåter information om den egna personen, som är fullständigare, direktare och konkretare än information via andra media.
3. Videobandade beteenden tillåter att man omedelbart repeterar inspelningarna (i princip ett oändligt antal gånger) för att kunna verifiera eller falsifiera information.

Ett antal mätinstrument har konstruerats för att kvantifiera individens förmåga (a) att tillägna sig information (percipiera), (b) att bearbeta information (värdera) och (c) att använda denna information (modifiera). Punkt c kan dock inom ramen för detta experiment enbart studeras som en förändring av individens referensram. Frågan i vilken utsträckning lkk har åstadkommit en modifiering av undervisningsbeteenden, skulle ha krävt åtminstone tre undervisningstillfällen för att man skulle ha kunnat studera utvecklingsförloppet. Syftet är alltså att mäta i vilken utsträckning en presentation av subjektiv och objektiv information enligt ovan påverkar lkk perception och värdering av det egna undervisningsbeteendet vid olika tidpunkter. Alla items i det för experimentet utvecklade skattnings- och värderingsinstrumentet (F III) refererar till "under denna lektion . . ." för att knyta lkk reaktioner till den aktuella situationen (episodiska bedömningar i motsats till dispositiva bedömningar).

Den situation som enligt definition betecknas som undervisning är definierad genom tre komponenter (1) edukator (person eller läromedelssystem), (2) edukand (person som skall påverkas) och (3) undervisningsintentioner (mål), ämnen, metodik (undervisningsförfarandet) och mediaval (kommunikationsmedel). Dessa tre objekt är definierade genom ett antal attribut (egenskaper, relationer). För att underlätta diskussionen presenteras i figur 12 nedan mätinstrumentens a priori-struktur inom ramen för en flerdimensionell modell. Mätningen avser de tre klassiska objekten (lärare, elev och stoff) som på X-axeln betecknas med (1) ego, (2) elev och (3) icke personella objekt (IPO).

På Y-axeln anges den andra dimension som inom experimentets ram definierar dessa tre objekt. Denna dimension beskriver ett kontinuum, vars ändpunkter är fysiska respektive kommunikativa aspekter. Personens första reaktion på den egna videobandade prestationen gäller för det mesta de fysiska



Figur 12. A priori-strukturmodell för mätinstrument

aspekterna. Allen & Ryan (1969, s 55) betecknar denna reaktion som "cosmetic effect". Begreppet avser individens fokusering på t ex frisyr, maner (tics), anatomiska egenheter (anomali) etc. Allen & Ryans erfarenheter är att denna fokusering är relativt kortvarig men att detta stadium är mycket ömtåligt. På grund av individens (lk) allmänna osäkerhet rörande sin första undervisningsprestation, bör fp under denna första kritiska period — dvs fokusering av perception och värdering på individens *externa* gestalt — få så mycket tid och så många uppspelningar, att han kan åstadkomma *kongruens* mellan interna och externa föreställningar, innan någon mera systematisk påverkan via t ex hl kan påbörjas.

Som tidigare nämndes visar ett antal undersökningar att individen i en självkonfrontationssituation ofta i början reagerar med en "extrem" emotionalitet och "rejection" samt med försök att undvika den information som förmedlas via ITV/VR. Trots denna i början emotionella reaktion tyder undersökningsresultat (se t ex Nielsen, 1962; Walz & Johnston, 1963) på att självkonfrontation efterhand har en välgörande (i betydelsen objektiverande) effekt på individen.

Hypotesen skulle kunna vara att det perceptuella (modifiering av externa tecken) och det emotionella (modifiering av interna tecken) försvaret antas övergå i en fokusering på kognitiva och slutligen kommunikativa aspekter. Denna indelningsprincip (fysiska, emotionella, kognitiva aspekter) baseras på Wundts (1918, s 31) välkända klassificerings- och relationsprincip, medan kommunikativa beteendeaspekter är en mera unik kategori som faller utanför denna ram. Men denna kategori är i detta sammanhang av intresse, eftersom den beskriver sådana beteendeaspekter, som hänför sig till "lärar-elev"-interaktionen, dvs överföring av information.

Det antas att lkk efter att ha explorerat apparitiva, emotionella och kognitiva aspekter av de olika objekten riktar sin uppmärksamhet på kommunikativa beteendeaspekter, vilket i detta sammanhang innebär en medveten, målinriktad selektion och transformation av information. Denna abstraktionsnivå förutsätter en framgångsrik självdistansering, en emotionell och kognitiv objektivering av eleverna och ämnet. Individen har i detta fall lyckats övervinna sina "försvarsattityder", dvs är inte längre ego-centrerad.

En optimal självdistansering beträffande de ovan angivna tre parametrarna borde kunna innebära att de skattade beteendena får en korrekt värdering (t ex att önskade kommunikativa beteendeaspekter får den mest positiva värderingen på den sjugradiga skattningsskalan). Individens fullt utvecklade och optimalt fungerande "själv" som "reglerings- och kontrollinstrument" kan hypotetiskt definieras.

Man skulle kunna ställa upp hypotesen att i samma takt som individens förmåga till självdistansering växer, ökar hans förmåga till en mera realistisk eller objektiv syn. Inom ramen för denna modell blir det möjligt att ange på vilket sätt lkk, beroende på experimentets påverkningsformer, percipierar och värderar sig själva och andra objekt, definierade genom de attribut, som ingår i mätinstrumenten.

8.2 Beskrivning av mätinstrumenten

För denna studie ansågs två olika slags data vara av speciellt intresse för att belysa lkk reaktion på experimentets påverkningsformer: (1) lkk obundna och spontana muntliga kommentarer och (2) lkk skattningar på en sjugradig skala av ett antal beteendeaspekter. I gul bilaga 4 (Bierschenk, 1972 e) ges en uppställning över alla mätinstrument som kom till användning.

Fördelar och nackdelar med att använda skattningsskalor är allmänt kända. Skalorna betraktas rent pragmatiskt som "lika intervall-steg"-skalor, eftersom det är mycket svårt att empiriskt bevisa att skalorna verkligen har de egenskaper som teoretiskt sett beskriver en intervallskala. Intervallskalorna har fördelen att ge information om avståndet mellan förankringspunkterna. Wood-

worth och Schlosberg (1954) betraktar en intervallskala som en idealisk bedömningsskala. Men Torgerson (1958) betvivlar möjligheten att subjektiva bedömningar skulle kunna resultera i lika skalsteg. Skattningsskalorna bör därför ses som en subjektiv estimeringsmetod. Guilford (1954) betraktar å andra sidan denna typ av skattningsskalor helt enkelt som en enkel och ekonomisk metod men poängterar de olika problem som metoden medför (jfr De Lemos, 1963).

8.2.1 Självexplorering och självvärdering (Formulär F I, F II, F III)

Skattnings- och värderingsinstrument (Formulär F III)

Formulär F III (Bierschenk, 1972 e, gul bilaga, 4:9—4:30) innehåller 79 variabler uppdelade i (a) en skala för skattning av förekomst av eller kvalitet hos ett visst attribut och (b) en skala för värdering av detta attribut i förhållande till den aktuella mikrolektionen. För varje skattning anges dessutom om lk är (a) ganska säker eller (b) mycket osäker. Formuleringen ”ganska säker” valdes i stället för mycket säker (helt säker) för att undvika att alltför många individer av försiktighetsskäl skulle välja alternativ b. Lkk utförde bedömningarna av de olika attributen enligt följande exempel.

Ruta 2. Exempel på variablerna i formulär F III

21 a) Under denna lektion talar jag — BEDÖMNINGEN är

	1 2 3 4 5 6 7		
utan dialekt alls		med mycket framträdande dialekt	a ganska säker b mycket osäker

b) Att tala med dialekt under denna lektion bedömer jag för egen del som

	1 2 3 4 5 6 7		
mycket besvärande		helt obesvärande	a ganska säker b mycket osäker

c) Att tala dialekt under denna lektion bedömer jag för elevernas del som

	1 2 3 4 5 6 7		
mycket besvärande		helt obesvärande	a ganska säker b mycket osäker

Tabell 12. Subjekt-objekt-relationer, a priori fördelning av påståendena i skattnings- och värderingsformulär F III

SUBJEKT	OBJEKT: 1. EGO	OBJEKT: 2. ELEV	OBJEKT: 3. ICKE-PERS. OBJEKT
EGO 1	1, 2, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 66, 71 Frågor med c-alternativ: 10, 14, 21	17, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 50, 51, 52, 56, 59, 72, 73, 74, 75, 78 Frågor med c-alternativ: 32, 34	3, 6, 61, 62, 63, 64, 65, 67, 68, 69, 70, 76 Frågor med c-alternativ: 6, 63, 70
ELEV 2	41, 42, 43, 44, 45, 46, 47, 54, 55, 58 Frågor med c-alternativ: 47, 55	40, 48, 49, 53	57, 60, 77, 79

Utttrycket "Under denna lektion" har använts i syfte att knyta lkk skattningar till episodiska bedömningar och undvika dispositiva bedömningar. Det stora flertalet av attributen har enbart alternativen a och b (jfr tabell 12). Endast påstående 6 hade fått ett d-alternativ, men detta togs ej med i bearbetningen, då skalan ej är bipolär. De 79 items som ingår i formulär F III beskriver de sex mätområden som tabell 12 visar, dvs de fördelar sig enligt modellen för ett studium av lkk *subjekt-objekt-relationer*. Vid utvärderingen har i första hand tagits hänsyn till dimensionerna a och b. Formulär F III administrerades omedelbart efter varje påverkan. Lkk fick före experimentets början ej någon övning eller kännedom om innehållet i formulär F III för att vi skulle kunna garantera, att åtminstone bedömningen efter första självkonfrontationsupplevelsen ej var styrd av experimentatorn. Skattningarna utfördes dessutom på grund av simultankommentarerna retrospektivt.

Formulär F III delades upp i tre huvudkategorier med varierande antal underkategorier, enligt uppställningen i ruta 3, som operationellt definierades av de angivna items:

Ruta 3. Kategorier och items i formulär F III

A	*Bedömning av min egen person*		
	I	Mina emotionella reaktioner	1–6
	II	Röst, röstläge	7–9
	III	Motorik	10–15
	IV	Kunskaper	16
	V	Uttrycksförmåga	17–23
B	*Bedömning av elevernas beteende mot mig och elevernas beteende sinsemellan*		
	I	Mitt sätt att leda klassen	24–35
	II	Min uppmärksamhet mot vissa elevtyper	36–37
	III	Min kontakt med eleverna	38
	IV	Disciplinära åtgärder	39–40
	V	Elevernas aktivitet riktad mot mig	41–47
	VI	Elevernas kontakt sinsemellan	48–49
	VII	Bedömning av elevernas fysisk-psykiska tillstånd	50–53
	VIII	Bedömning av elevernas intellektuella aktivitet	54–58
C	*Min undervisningsplanering m m*		
	I	Bedömning av förutsättningarna för undervisningsplanering	59–60
	II	Planeringens struktur	61–62
	III	Hjälpmedel	63
	IV	Att använda tavlan	64–66
	V	Uppföljning av metodiska steg	67–70
	VI	Mitt sätt att ställa frågor	71–75
	VII	Buller och oväsen utifrån	76–78
	VIII	Studio-situationens påverkan på eleverna	79

Skattnings- och värderingsinstrumentets kortare version (Formulär F I)

Formulär F I (Bierschenk, 1972 e, gul bilaga 4:1—4:5) innehåller 22 variabler som till innehåll och formulering är ekvivalenta med formulär F III. En identisk formulering av påståendena hade varit önskvärd, men eftersom formuleringarna i formulär F I skulle referera till lkk upplevelse av undervisningsförloppet och formuleringarna i formulär F III till lkk bedömning av undervisningsförloppet, skiljer sig dessa påståenden i formuleringen, men är identiska i fråga om innebörden.

Ur varje underkategori i formulär F III valdes ett item. I tabell 13 ges en uppställning av de utvalda items:

Tabell 13. Motsvarande item i F I och F III

F I	F III	F I	F III
1 a	6 a	11	39 a
1 b	6 b	12	41 a
2 a	79 a	13	48 a
2 b	79 b	14	50 a
3	1 a	15	50/58
4	9 a	16	59 a
5	11 a	17	61/62
6	15 a	18	64 a
7	17 a	19	63 b
8	24 a	20	67 a
9	36 a	21	76 a
10	38 a	22	77 a

Formulär F I utdelades efter andra undervisningstillfället och före den efterföljande uppspelningen. Syftet med detta instrument är att studera eventuella skillnader i perception av den omedelbara undervisningsupplevelsen och den efterföljande självkonfrontationsupplevelsen. För att något minska belastningen för lkk har formulär F I utformats som skattningsformulär. Skattningarna utfördes enligt exemplet i ruta 4:

Ruta 4. Exempel på variablerna i formulär F I

1. Under denna lektion kände jag mig

mycket passiv	1	2	3	4	5	6	7	mycket aktiv

BEDÖMNINGEN är
a ganska säker
b mycket osäker

Identifikationsupplevelse (Formulär F II)

Formulär F II (Bierschenk, 1972 e, gul bilaga 4:6—4:8) konstruerades i syfte att undersöka på vilket sätt lk upplever den egna personen i självkonfrontationssituationen med hänsyn till (a) sättet att percipiera och (b) upplevelsen av denna situation (11 frågor) samt (c) lk önskan att se uppspelningen av de videobandade mikrolektionerna ensam eller tillsammans med någon annan person (fråga 12). Skalorna utformades enligt exemplet i ruta 5:

Ruta 5. Exempel på variablerna i formulär F II

Under denna uppspelning är min uppfattning om eleverna

1	2	3	4	5	6	7
helt förändrad						helt oförändrad

I fråga 12 fick lkk rangordna 8 möjliga alternativa former som t ex att vara ensam under självkonfrontationsprocessen, att vara tillsammans med någon annan person eller att beskriva något önskat icke listat arrangemang.

Formuleringarna i de ovan presenterade tre skattningsschemata har utformats så, att skattningarna är direkt relaterade till den givna situationen. Avsikten med den permanenta anknytningen till själva undervisningssituationen har varit att komma åt den situationsbundna upplevelsen och ej lkk uppfattning om undervisningssituationer i största allmänhet. Även i den mån ett visst beteende ej förekom under undervisningens gång, fick lkk under (a) ange att det ej har förekommit och under (b) ändå värdera den icke uppkomna situationen som om den hade förekommit "i måttlig utsträckning". Det visade sig vid en kontroll att lkk av sig själva reagerade på ett sådant sätt, utan att detta explicit hade meddelats i anvisningarna. Ett annat alternativ hade ju varit att indikera att ett beteende är obesvärande, när det inte förekommer, men det hade varit en tämligen trivial information.

8.2.2 Elevattityder (Formulär F IV)

Trots att det ej ingår i experimentets målsättning att mäta dess effekter på elevgrupperna, har vi dock velat få en viss uppfattning om på vilket sätt eleverna upplever deltagandet i undersökningen. Omedelbart efter varje undervisningstillfälle fick eleverna besvara ett attitydformulär (F IV omfattande 8 frågor som redovisas i Bierschenk, 1972 e, gul bilaga 4:31—4:32) enligt exemplet i ruta 6:

Ruta 6. Exempel på variablerna i formulär F IV

1. Tycker Du det var spännande i TV-rummet?	Ja	Nej	Vet ej

8.2.3 Forskningsvärdering (Formulär F V—F VII)

Formulär F V (Bierschenk, 1972 e, gul bilaga 4:33—4:35) konstruerades för att vi skulle få en uppfattning om på vilket sätt de lkk som deltog i experimentet upplevde denna försöksverksamhet. För att få ett visst mått på experimentets yttre validitet, konstruerade vi ett attitydformulär (omfattande 10 frågor) för de lkk som deltagit i ITV-försöket. Instrumentet betecknas: *Forskningsvärdering (deltagare)*.

Vi vill även få en uppfattning om hur de ej deltagande lkk upplevde ITV-försöket. Detta instrument (F VI) betecknas: *Forskningsvärdering (ej deltagare)*. (F VI, omfattande 6 frågor, redovisas i Bierschenk, 1972 e, gul bilaga 4:36—4:37.)

Genomförandet av experimentet medförde, att även annan undervisning blev berörd av försöksverksamheten. I syfte att ge en uppfattning om attityden hos inom LHM undervisande personal konstruerades instrumentet F VII omfattande 10 frågor, redovisade i Bierschenk, 1972 e, gul bilaga 4:38—4:40, och betecknas: *Forskningsvärdering (personal)*. Resultaten av attitydmätningen med hänsyn till F V, F VI och F VII redovisas i kapitel 12.

8.2.4 Simultankommentarer

Att använda skattningsskalor kan ha olika nackdelar, bl a att man kanske går miste om för individen mycket väsentliga dimensioner, dvs att man på ett icke önskat sätt begränsar fpp svarsmöjligheter. För att kunna komplettera skattningsskalorna och för att bereda fpp tillfälle att ge verbala responser efter fritt val, fick fpp möjlighet att yttra sig helt spontant och obundet under de olika experimentella betingelserna. Arrangemanget för simultankommentarerna fanns vid alla uppspelningstillfällena. Instruktionen till fpp i samband med simultankommentarerna var att kommentera sig själva och undervisningsskeendet samtidigt (simultant) med uppspelningen. För detta syfte hade vi en anordning av instrument där fpp via hörlurar kunde följa undervisningen. Via en stativ- respektive bordsmikrofon överfördes fpp kommentarer till en ljudbandspelare, som var uppställd i kontrollrummet. Simultankommentarerna blev mycket varierande under förförsöket. För att ändå få en viss information av de fpp som eventuellt ej skulle säga någonting, fick alla fpp ge några skriftliga synpunkter.

8.2.5 Öppna formulär (Formulär FA, FB, FC)

Omedelbart efter avslutad första mikrolektion uppmanades fpp att under ca 5 min skriva ner hur de upplevde denna. För att inte styra informationen fick fpp vid detta tillfälle formulär FA (Bierschenk, 1972 e, gul bilaga 5:1) med följande uppmaning: "Du har just avslutat Din lektion. Hur upplevde Du *Dig själv, eleverna, undervisningen* och *TV-studion?* Skriv ner Dina upplevelser. Du har fem minuter på Dig."

I anslutning till experimentet började lkk sin praktikvecka. Detta anges i figur 11 med Pr. För att visa eventuella effekter av experimentets påverkningsformer på lkk perception och värdering av undervisningsupplevelserna under praktikveckan samt för att få en uppfattning om fpp och de övriga lkk (som ej deltog i experimentet) reaktioner konstruerades formulär FB (Bierschenk, 1972 e, gul bilaga 5:2—5:4). Detta mätinstrument består av sju öppna frågor. Sex motsvarar de sex fälten i subjekt-objekt-relationsparadigmet. Den sjunde frågan berör upplevelsen av handledningen under praktikveckan.

Vid slutet av försöksverksamheten, dvs vid slutet av andra terminen, var vi intresserade av att få några synpunkter från lkk på frågor, som aktualiserades under experimentets gång. Formulär FC (se Bierschenk, 1972 e, gul bilaga 5:5—5:8) innehåller ett antal sådana frågor samt en uppmaning i form av en öppen fråga att sammanfattningsvis värdera hela experimentet.

8.2.6 Lärarkandidaters predispositioner och personlighetens inflytande på lärarkandidatens perception och värdering av videobandförmedlade undervisningsprocesser

Undersökningar som avser informationsselektion och beteendemodifikationer till följd av individens predispositioner (se kap 3.2 ovan) utgår från det allmänna antagandet att individens personlighet är av central betydelse för arten och omfattningen av beteendemodifikationer. Ett resultat i samband med undersökningar angående externt förmedlad självkonfrontation via ITV/VR är bl a att fpp belåtenhet med den egna prestationen före uppspelningen kommer att bestämma, vad som observeras på en TV-monitor, på vilket sätt detta värderas och till vilka attitydförändringar det leder (se Salomon & MacDonald, 1969, s 15). Sådana resultat implicerar att man också borde undersöka sambandet mellan de experimentella resultaten och olika personlighetsvariabler. De test som presenteras nedan har utvalts för att eventuellt ge oss möjlighet att studera, i vilken utsträckning lkk perception och värdering av den egna undervisningen styrs eller påverkas av individens (1) kognitiva förmåga, (2) förmåga att hålla emotionell balans, (3) tillgång till adekvata och sociala beteenden, (4) tillgång till ett elevanpassat (konkret och enkelt) språk, (5) förmåga att stimulera och kontrollera undervisningsprocessen, (6) förmåga att motstå åsiktsförändringar under inflytande av olika provokationstyper, (7) förmåga till ett integrativt beteende, (8) förmåga att acceptera sig själv och andra, (9) förmåga till perceptuell analys och (10) tillgång till en hög energi- och uppmärksamhetsnivå etc.

Det finns naturligtvis många olika tillvägagångssätt att skatta dessa egenskaper. I samband med intagningen av studerande till klasslärarlinjen vid LHM har i flera år i forskningssyfte administrerats ett grupptest-batteri som innehåller sådana test som avser att mäta dessa variabler. Detta testbatteri har varit en naturlig utgångspunkt i ett första försök att kartlägga i vilken utsträckning lkk perceptions- och värderingstendenser under experimentet är beroende av sådana dispositiva egenskaper som angivits ovan. En mera ingående diskussion av de olika testen och en presentation av den teoretiska utgångspositionen för urvalet av dessa test finns i Bjerstedt & Sundgren (1968) och Bjerstedt (1967 c). Den totala testtiden för varje studerande har varit ungefär sex timmar. Testen administrerades i två eller tre testomgångar omfattande två eller tre dagar. Med hänsyn till de ovan presenterade 10 problemställningarna valdes ur detta testbatteri nedanstående test. Resultat från dessa test kommer emellertid ej att redovisas i detta sammanhang.

Beteckning	Kort beskrivning
1. WIT III	är ett standardiserat svenskt intelligenstest, utvecklat av Westrin (1964). Det innehåller fem deltest: (1) Klassifikationer, (2) Analogier, (3) Motsatser, (4) Sifferkombinationer och (5) Pussel.
2. Cattell 3:A	är det välkända "IPAT Culture Fair Intelligence Test", Skala 3, Form A, utvecklat av Cattell & Cattell. Det innehåller fyra icke verbala deltest: (1) Serier, (2) Klassifikationer, (3) Matriser och (4) Betingelser. (En svensk upplaga har sammanställts efter överenskommelse mellan IPAT och Bjerstedt.)
3. Rosenzweig Vuxenversion	är det välkända "bild-frustrationstestet", utformat av Saul Rosenzweig. Testpersonen har till uppgift att ta ställning till 24 illustrerade frustrationssituationer. Den svenska adaptationen beskrivs av Bjerstedt (1968 a).
4. Språkförenkling Version C	avser att mäta fpp förmåga att transformera komplexa och abstrakta uttryck till enklare språkformer. Testet består av 51 items. (Icke publicerat.)
5. Skolsituationer	bygger på Rosenzweigs metodik. Testet består av 10 mer eller mindre vardagliga skolsituationer. (Icke publicerat.)
6. Åtgärder vid disciplinproblem	avser att mäta fpp benägenhet att använda olika grader av tillrättavisande av elevbeteenden. Testet innehåller 42 olika elevbeteenden och elevreaktioner. (Icke publicerat.)

7. Lärarreaktioner	simulatortest som finns i tre versioner. Varje version innehåller fem videobandade scener. Varje scen behandlas ur olika aspekter. Svaren avges dels i bunden, dels i öppen form. Testet beskrivs av Bjerstedt (1968 b).
8. Formulär EK 9	är ett urval av nio skalor ur EPPS (Edwards Personal Preference Schedule). EK 9 avser att mäta 9 personlighetsvariabler (a–i) och en kontrollvariabel (j). I testet ingår (a) Achievement (ach), (b) Deference (def), (c) Order (ord), (d) Autonomy (aut), (e) Affiliation (aff), (f) Intraception (int), (g) Nurturance (nurt), (h) Change (chg), (i) Aggression (agg), (j) Consistency (con). (Testet har adapterats av Bjerstedt.)
9. Nutidsfrågor Version C	är ett multiple-choice test, konstruerat med syfte att indikera lkk tendens att uppleva hot, isolering och aggressivitet i skol- och allmänna situationer, parti för eller emot "tredje ståndpunkten", "mänskliga rättigheter" och "världsmedborgarsinne" samt "läroplansorientering". Testet innehåller 74 påståenden som indikerar sju skalor.

Fullständigheten av dessa testdata avseende de lkk som skulle delta i experimentet, kontrollerades under lk första termin. I de fall testformulär saknades, genomfördes eftertestningar före experimentets början.

Det ursprungliga testbatteriet, som har beskrivits av Bjerstedt & Sundgren (1968), utformades år 1962, men det har efterhand undergått åtskilliga modifikationer (test har utgått och andra har tillkommit). Testbatteriet som beskrivs nedan innehåller dels test som tidigare har använts i samband med intagningstestningarna (nr 11, 12, 13, 14), dels vissa nya test vilka anses vara relevanta variabler i samband med experimentet (nr 15, 16, 17, 18).

Externt förmedlad självkonfrontation via ITV/VR (se kap 7) antas medföra en de-automatiseringsprocess av det vanliga sättet att uppleva sig själv. Det antas vidare, att detta tillstånd möjliggör en förändring i en relativ autonomi, antingen i form av en ökning eller minskning. Processen innebär således dels en "omskakning" som kan följas upp genom en förbättring eller en regress i individens organisationsnivå, dels en re-dirigering av uppmärksamheten på sådana processer, för vilka den ej längre varit nödvändig, eftersom funktionerna hade automatiserats.

Dessa antaganden medför att, utöver de testvariabler som har presenterats ovan, samband mellan lkk perceptions- och värderingstendenser och de variabler som presenteras nedan kan bli av intresse i undersökningen.

Heil, Powell & Feifers (1960) undersökning, som refererades inledningsvis, visade att lärarens personlighet är av stor betydelse för lyckat arbete med olika elevtyper. I socialpsykologiska studier har man kunnat konstatera samband mellan individens självaccepterande attityder och accepterande av andra personer eller persongrupper (Berger, 1952, s 778). Sådana resultat skulle

kunna implicera att personer som ej accepterar eller har svårt att acceptera sig själva är fientligt inställda mot andra personer eller persongrupper. Att närmare kunna granska relationen mellan eventuella specifika perceptions- och värderingstendenser under experimentet och accepterandet av sig själv eller andra skulle kunna underlätta klargörandet av problem i samband med sociala konflikter.

Vad fpp kommer att observera på en TV-monitor skulle t ex kunna antas vara beroende av ett "rädsla-isolation-aggressions"-syndrom som definierats av Bjerstedt & Sundgren (1968, ss 57—59). Samband mellan lkk perceptions- och värderingstendenser och detta syndrom skulle kunna undersökas med hjälp av t ex test nr 3, 5, 6, 8, 9 ovan och nr 13 nedan. Även lkk attityder mot olika skolaktiviteter och -situationer kan tänkas ha samband med lkk sätt att percipiera och värdera den egna, via ITV/VR förmedlade, undervisningssituationen.

Den framtida läraren förväntas visa interaktionsbeteenden som överensstämmer med läroplanens rekommendationer, men eftersom dessa i en pågående skolreformverksamhet kontinuerligt förändras, krävs det lärarpersonligheter som utmärkes av bl a dels flexibilitet, dels autonomi. Relaterat till denna problemställning skulle det kunna vara av intresse att studera eventuella samband mellan lkk perception och värdering av de egna mikrolektionerna och dels lkk förändringsbenägenhet (flexibilitet), dels deras förmåga att motstå prestigeladdade förändringsprovocerande kommentarer.

De test som skulle kunna användas för att undersöka sådana eventuella samband och som ingick i det testbatteri, som administrerades vid slutet av lkk andra lärarhögskoletermin, är följande:

Beteckning	Kort beskrivning
10. USTM-provet	är ett intelligenstest av grupptesttyp. Det består av fyra delar. Uppmärksamhet, Synonymer, Talserier och Matriser. Testet har utvecklats av Husén & Henricsson (1950). Den aktuella manualen till USTM-provet utformades av Prien. Provet är avsett att ge en nivåbestämning av intelligensen. I samband med experimentet administrerades enbart delprov A: Uppmärksamhet. Detta delprov innehåller 43 uppgifter. Testet är inte faktorrent utan mäter närmast en allmän g-faktor. Deltest A avser allmän uppmärksamhet.
11. Preferenser i skolsituationen	är en svensk adaptation av Sterns "Teacher Preference Schedule, Form G". Det innehåller tio skalor för att mäta olika lärarroller. Testet är beskrivet i Sundgren (1966).
12. Dolda figurer	är ett test av Gottschaldt-typ. Testet är konstruerat av Cattell (test 37, se Cattell & Warburton, 1967). Testpersonen har till uppgift att lokalisera enkla geometriska former inom ramen för mera komplexa figurer. Testet är strikt tidsbegränsat. Det avser att mäta perceptuell analytisk förmåga ("fältartikulering"). (En svensk upplaga har sammanställts av Bjerstedt i samråd med IPAT.)

13. 16 PF – Form B är en svensk adaptation av Cattells 16 PF (The Sixteen Personality Factor Questionnaire). Testet är avsett att mäta 16 faktorer med hjälp av 186 påståenden. Testadaptationen har utförts av Bjerstedt.

14. Personliga åsikter (PÅ) är en svensk adaptation av ett Cattell-test, som består av två delar: "Block X Composite" resp. "Block Y Composite", deltest i Cattells "O-A Personality Test Battery". På motsvarande sätt finns det för den svenska versionen PÅ Del I och PÅ Del II. Den längre versionen innehåller 100 åsiktspåståenden, den kortare versionen (som har använts i experimentet) endast 50. Lk anger sin inställning till dessa på en femgradig skala. Syftet är bl a att studera individens tendens att modifiera åsikter under inflytande av olika typer av inbyggda förändringsprovokationer. Påståendena som Del I innehåller är i Del II upprepade och försedda med förändringsprovocerande kommentarer, prestige-laddade ursprungsbeteckningar eller kvasi-uppgifter från föregivna opinionsundersökningar. Adaptionen har utförts av Bjerstedt (1965).

15. R 10 F är en svensk adaptation av "the California F Scale" (Adorno m fl, 1950). Skalan består av 25 påståenden och är avsedd att mäta allmänna auktoritära tendenser. Adaptationen har utförts av Rubenowitz (1963).

16. R 10 D är en svensk adaptation av Rokeach's "Dogmatism Scale" (Rokeach, 1960). Skalan består av 26 påståenden och är avsedd att mäta allmänna dogmatiska tendenser. Adaptationen har utförts av Rubenowitz (1963).

17. Formulär F VIII är en svensk adaptation av Bergers test "The relation between expressed acceptance of self and expressed acceptance of others". Testet innehåller två skalor (1) själv-accepterande, indikerad med 36 items och (2) accepterande av andra, indikerad med 28 items. Testet är beskrivet av Berger, 1952. Adaptationen har utförts av Bierschenk.

18. Formulär F IX a, b, c, d. är en nykonstruktion. Testet är avsett att mäta individens (a) allmänna självvärdering, självvärdering i förhållande till (b) åtta konkret beskrivna undervisningssituationer, till (c) åtta konkret beskrivna studiesituationer och till (d) åtta konkret beskrivna familjesituationer. (Testet har konstruerats av Bierschenk.)

9. Datauppläggning och datakontroll

Som framgår av föregående kapitel har mätningarna i samband med experimentet resulterat i en stor datamängd. Det relativt stora antalet mätinstrument är beroende av experimentets explorativa karaktär. Många av dessa mätinstrument är nykonstruktioner och kräver en speciell behandling i samband med analysarbetet. Vilka data som blir föremål för en mera ingående analys är beroende av de problem som uppstår under bearbetningens gång. Den primära analysen koncentreras dock kring huvudinstrumentet (Formulär F III).

En första databehandlingsuppgift är att strukturera, organisera och lagra den stora datamängden på ett sådant sätt, att den är lätt åtkomlig för alla de olika analysproblem som kan uppstå. För att kunna analysera experimentets data i olika etapper krävs en datamaskinell behandling. Vid bearbetningen av en större datamängd måste denna ha överförts till register som under analysarbetets gång dels levererar den önskade informationen, dels lätt kan uppdateras eller omstruktureras.

Omedelbart efter slutförandet av respektive experimentell fas påbörjades kodnings- och stansningsarbetet, dvs överföring av datamängden på IBM-hålkort för senare behandling vid LDC (Lunds Datacentral). Stansningen av hela datamängden utfördes av institutionens stansoperatris. Men eftersom institutionen enbart är utrustad med en vanlig hålkortsstans överlämnades ett urval av materialet för kontrollstansning till IBM-företaget i Malmö.

Materialet för kontrollstansning bestämdes enligt följande förfaringssätt. Först bestämdes att ca 10 % av materialet skulle vara tillräckligt för upptäckande av fel och felens karaktär. Av 48 lkk för respektive fas utvaldes slumpmässigt sex lkk, vilkas hela datauppsättning skulle kontrollstansas. Även med vilken lk uttagningen skulle börja bestämdes slumpmässigt. Därefter uttogs var åttonde lk för att täcka över hela stansningstiden. På så sätt utfördes kontrollstansningen enligt samma mönster som lkk har gått genom försöket och som samtidigt har varit ordningen för den ursprungliga stansningen. Även bedömningspanelens material kontrollstansades enligt samma förfaringssätt.

Vilken felmarginal som kan tolereras blir nödvändigtvis föremål för ett subjektivt avgörande. Vår kontrollstansning resulterade i en felmarginal på 2 ‰—3 ‰, som i detta sammanhang kan anses vara så obetydlig, att en kontrollstansning av hela materialet ej ansågs vara nödvändig. Till detta kom en kontroll av datamängden med hänsyn till logiska och systematiska fel. Denna kontroll skedde dock först efter överföringen till magnetband. Ett antal speciella program konstruerades för detta syfte.

För databehandlingen överlämnades hålkorten till LDC för maskinell behandling. Ett första steg var här att överföra data från hålkorten till magnetband. LDC utnyttjade fram till nyåret 1970 Uppsala Datacentrals maskin CD 3600 tills dess att en Univac 1108 hade installerats i Lund. CD 3600 matades från Lund med hjälp av en medelhastighetsterminal CD 8090 via telenätet. Datamaterialet från år 1969 har delvis bearbetats på CD 3600 och fanns där på ett antal magnetband i Uppsala. Dessa band konverterades våren 1970 i Lund för att passa Univac 1108.

Olika delar av datamaterialet skulle i stor utsträckning behandlas var för sig. Ett antal motsvarande dataposter byggdes upp på olika band för att man inte skulle behöva använda band med för en viss analys delvis irrelevanta data. Den interna representationen av data på band gjordes sådan, att dessa direkt kunde läsas från t ex BMD-programmen (en samling statistiska standardprogram utgiven av Dixon, 1970) och specialprogram skrivna på programmeringsspråken FORTRAN och ALGOL.

En ofullständig datauppsättning medför vanligen vissa begränsningar i databearbetningen och utnyttjandet av t ex BMD-programmen. För kontroll av hela datauppsättningens utseende, dvs eventuellt bortfall och fördelningar, har hela materialet beskrivits med hjälp av BMD01D. Bortfallet av såväl lkk som bortfall i svar eller markeringar i de enskilda mätinstrumenten är så obetydligt (ca 2 %) att datauppsättningen kan betraktas som fullständig. I de fall där svaret saknades på en viss fråga och en metod (t ex faktoranalys) krävde ett fullständigt material, genomfördes en medelvärdesberäkning för respektive variabel och närmaste heltal fördes in i luckan. I de fall där lk missade ett helt instrument, leder korrektionen till en neutralisering av individens värden. Men eftersom en sådan korrektion endast blev nödvändig i två fall och ett enda test, kan denna korrektion betraktas som utan effekt på materialet i sin helhet. När det i fortsättningen hänvisas till körningar av BMD-programmen, så har dessa ofta måst föregås av speciell programmering för att ur de fullständiga registren särskilda urval och omformning av data skulle kunna göras.

10. Databearbetning

Experimentets datamängd kan delas upp i två grupper, sådana data (simultankommentarer, öppna frågor) som kräver någon form av innehållsanalytisk bearbetning och sådana data som baseras på skattningsskalor och därmed mera direkt lämpar sig för parametriska analystekniker. I fortsättningen gäller diskussionen enbart de data som baseras på skattningsskalorna, dvs experimentets numeriska data. För diskussion av den innehållsanalytiska bearbetningen, se Bierschenk (1972 f).

Föremål för analysarbetet blev i första hand formulär F III, eftersom det för detta instrument föreligger data för den fullständiga designen. Som tidigare anförts har vi konstruerat formulär F III med utgångspunkt från subjekt-objekt-relationsmodellen. Alla andra formulär är avsedda att mäta vissa delproblem. Nedan anges de analyser som vi har utfört för att belysa de tre första i kapitel 2 nämnda problemområdena.

För att kunna analysera datamängden i enlighet med experimentets problemställningar blir såväl nivåanalyser som struktur- eller dimensionsanalyser nödvändiga. Dessutom kan hela datauppsättningen delas upp i fem större delmängder: (1) lk självskattningar som resultat av experimentets olika påverkningsformer, (2) data som resultat av bedömningspanelens skattningar, (3) en kombination av data som resultat av lkk och bedömningspanelens skattningar och (4) bakgrundsdata som resultat av kognitiva test, personlighetstest och attitydtest. Därtill kommer (5) "kringdata" i form av t ex data om elevreaktioner, personalsynpunkter o dyl.

10.1 Lärarkandidaters självbedömning

1. *Nivåanalys* (utförd med BMD08V)

I syfte att undersöka om det finns signifikanta huvud- eller samspelseffekter genomfördes sex variansanalyser (en för varje formulärområde i F III). För att lösa vissa mättekniska problem bildade mätinstrumentet en faktor med ett varierande antal nivåer beroende på antalet items i varje fält (jfr tab 12).

2. *Dimensionsanalys* (utförd med BMD03M)

Formulär F III är en nykonstruktion, som utvecklades speciellt för detta experiment. Som ovan diskuterades har uppdelningen av items på de sex olika fälten skett a priori. För att man skall kunna studera strukturen i instrumentet på mera objektiva grunder krävs det någon form av dimensionsanalys. Denna analys genomfördes för formulär F III med faktoranalysteknik och resulterade i åtta faktoranalyser. För att möjliggöra ett studium av vilka items som konsekvent är representerade i samma faktorer följer dimensionsanalysen samma mönster som nivåanalysen:

1. Analyserna 1 och 2 är övergripande analyser, dvs för hela experimentella uppläggningen genomfördes en faktoranalys. Den första analysen utfördes med kriteriet egenvärdet 1.0 som gräns för faktorextraheringen och den andra analysen med skredtestet som kriterium (se t ex Bierschenk, 1971 b, s 43). Granskningen av egenvärdesfördelningen i den första analysen ledde till att de sju första faktorerna roterades. De övriga 6 analyserna är delanalyser.
2. Analys 3 avser ITV/VR-betingelsen, dvs fält 1 och 3 sett över båda undervisningstillfällena (ml_1 och ml_2).
3. Analys 4 avser betingelser utan ITV/VR, dvs fält 2 och 4 sett över ml_1 och ml_2.
4. Analys 5 avser handledning, dvs fält 1 och 2, sett över ml_1 och ml_2.
5. Analys 6 avser utan handledning, dvs fält 3 och 4 sett över ml_1 och ml_2.
6. Analys 7 avser ml_1, sett över båda experimentella faktorerna.
7. Analys 8 avser ml_2, sett över båda experimentella faktorerna.

Detta analysförfarande tillåter att man mera objektivt (men fortfarande ej tillfredsställande) kan utkristallisera vilka items som är konstanta och samlade i en och samma faktor i samband med delanalyserna och den övergripande analysen. En sådan analys förutsätter en transformering av den tredimensionella dataparallellepipeden till en tvådimensionell datamatris som sedan tillåter utnyttjandet av tillgängliga faktoranalystekniker. För diskussion av detta förfaringssätt hänvisas till Horst (1965, ss 315—332).

3. *Reliabilitet*

Viss information om itemreliabiliteterna kan erhållas från kommunaliteten, som här skattas med en kvadrerad multipel korrelation R_i^2, beräknad genom att varje item successivt betraktas som beroende variabel med resterande items som oberoende variabler. Om antalet individer är stort i förhållande till antalet variabler, är R_i^2 en undre gräns för den klassiska reliabiliteten. I föreliggande fall kan möjligen R_i^2 användas i detta syfte i analyserna 1 och 2 ovan.

10.2 Expertbedömningar genom pedagogiklektorer

1. *Nivåanalys* (utförd med BMD08V)

Syftet med denna analys är att granska om det finns signifikanta medelvärdesskillnader mellan de två huvudbedömarna, dvs nr 1 och nr 2.

2 a. *Korrelationsberäkningar* (utförda med speciellt skrivna program, programmerare Leif Robertsson). Korrelationsberäkningar har dels genomförts med hänsyn till alla fyra bedömarna, dels med hänsyn till olika bedömningstillfällen för bedömare nr 1 och nr 2. Beräkningarna avser dels skattningen av en enskild individ sedd över alla items i formulär F III, dels skattningen av ett item sett över alla individer. Korrelationsberäkningar föreligger också för de s k principbedömningarna.

2 b. Korrelationsberäkningar för bedömare nr 1 och nr 2: *medelbedömning* (utförda med BMD08V). För att kunna få ett reliabilitetsmått på medelbedömningarna genomfördes för varje item en ANOVA. För diskussion av denna bearbetningsmetod hänvisas till Winer, 1962, ss 124—132 och till Medley & Mitzel, 1963, ss 310—320.

10.3 Jämförelse mellan expertbedömningar och självbedömningar

För att kunna studera och analysera skillnader och likheter mellan självbedömningarna och expertbedömningarna utfördes följande analyser.

1. *Nivåanalys* (utförd med BMD08V). I denna ANOVA studeras skillnaderna mellan självbedömning och bedömare nr 1 och nr 2 medelbedömning.

2. *Dimensionsanalys* (utförd med BMD06V). Denna är en kanonisk korrelationsanalys, avsedd att visa (a) om båda datauppsättningarna är signifikant relaterade till varandra och (b) på vilket sätt dessa variabeluppsättningar kan kombineras med varandra så att korrelationen mellan komponenterna blir ett maximum. För diskussion av denna metod hänvisas till Hotelling, 1935, ss 139—142.

En sammanfattning av de utförda analysprogrammen ges i ruta 6. En redovisning av respektive resultatanalys sker i delarna 2, 3 och 4.

Ruta 6. Sammanfattning av analysprogrammen

Statistik	Ikk Självbed	Bedömningspanel Expertbed	Ikk Självbed	+Bedömningspanel +Expertbed
Beskrivande statistik	BMD01D	BMD01D	BMD01D	
Reliabilitetsbedömning	Faktoranalys BMD03M 8 analyser	Eget program: Prod-mom-korrel (r) samt för medelbed BMD08V 158 analyser		
Nivåanalys	ANOVA BMD08V 6 analyser	ANOVA BMD08V 6 analyser	ANOVA BMD08V 6 analyser	
Dimensionsanalys	Faktoranalys BMD03M 8 analyser		Kanonisk korrelationsanalys BMD06M 24 analyser	

11. Experimentets inre och yttre validitet

11.1 Inre validitet

Kontrollerade experiment inom det pedagogisk-psykologiska området i samband med användningen av de i dag mera välkända statistiska analysteknikerna har en bara 25-årig tradition (se Stanley, 1967, s 2). Men kontrollerade experiment är den främsta metoden för att avgöra skilda uppfattningar om mest adekvat pedagogisk-psykologisk praxis. Coombs (1968, s 130) skriver:

"Visserligen har nyheter kommit till och det gamla förbättrats. Men för det mesta har de avlagrats som geologiska skikt på den grund av metodik och pedagogik som rått i generationer. Det har inte skett någon grundläggande omskakning och omgruppering av tillgängliga teknologier för att skapa en ny syntes av undervisningsmetoder... Så är fallet när t ex film eller arbetsböcker eller TV-lektioner eller språklaborationer helt enkelt läggs till vad som redan pågår."

Kontrollerade experiment är även i sådana fall den överlägsna metoden för att verifiera eller falsifiera olika slags förbättringar och för införandet av nya medier och metoder utan risk för ett ogrundat, svärmiskt avskaffande av gammal lärdom till föremål för underlägsna nyheter (jfr. Campbell & Stanley, 1963, s 172).

För att resultaten från ett kontrollerat experiment skall kunna tolkas och generaliseras måste undersökningen ha dels *intern,* dels *extern* validitet. Intern validitet är en grundförutsättning för att experimentet över huvud taget skall kunna tolkas, medan frågan om resultatens generaliserbarhet, dvs frågan om ett experiments externa validitet, aldrig kan besvaras fullständigt. Med hänsyn till den ständigt växande floden av pedagogiska forskningsresultat måste det vara av speciellt intresse att uppmärksamma undersökningarnas interna validitet för att utkristallisera användbara resultat. Att granska och sammanfatta enskilda forskningsresultat är ett mycket viktigt steg i generaliseringsprocessen. Bedömningen av hållbarheten, dvs tolkbarheten i enskilda resultat har inom ITV-forskningen lett till mycket olika omdömen.

Schramm, ledare av *Institute for Communication Research of Stanford* publicerade år 1962 en omfattande rapport med titeln *Educational Television: The Next Ten Years.* I denna rapport skriver han att han undersökt 393 forskningsresultat angående ETV och ITV. Han påstår på s 53:

"... we have been able to assemble 393 cases in which instructional television has been compared with other classroom teaching, 32 cases in which home instruction by television has been compared with classroom teaching, and 14 cases in which military instruction by television has been compared with face-to-face teaching, with what seemed to be adequate design, controls, and statistics."

Hans analys resulterar i påståendet att TV-undervisning för det mesta var överlägsen konventionell undervisning. Men vad menar Schramm med bestämningen "adekvat"?

Stickells (1963) analys av forskningsrapporter angående TV-undervisning baseras på 250 jämförelser. Hans resultat visar med hänsyn till undersökningarnas inre och yttre validitet, att 87 % av jämförelserna måste bedömas som "icke tolkbara" (Stickell, 1963, s 67). Han granskade resultatens tolkbarhet mot bakgrund av följande kriterier:

1. Urvalet av fpp måste med hänsyn till påverkningsformen vara slumpmässigt.
2. Experimentgrupps- och kontrollgruppsindivider måste ha blivit undervisade av samma lärare.
3. Antalet fpp i experiment- och kontrollgruppen måste åtminstone vara 25.
4. Kontrollgruppen måste uppfylla vissa minimikrav (jfr Stickell, 1963, s. 34).
5. De statistiska antaganden, som den använda analystekniken förutsätter, måste ha blivit testade och visat sig hållbara.
6. Det måste ha kartlagts, att det ej fanns några andra viktiga systematiska skillnader utöver den experimentella påverkningen mellan experiment- och kontrollgrupp.
7. Mätinstrumenten måste ha blivit granskade i fråga om validitet och reliabilitet och bedömda som tillfredsställande.

Stickells undersökning visar, att endast 4 % av undersökningarna utnyttjar kontrollgruppsdesignen med "tillfredsställande" kontrollgrupper, att åtminstone 17 % av undersökningarna baseras på slumpmässigt utvalda grupper och att ca 50 % av undersökningarna har utförts utan någon experimentell kontroll (Stickell, 1963, s 46) över urvalet av medverkande individer. Allt som allt visar Stickells granskning (s 48) dels att endast 10 av de 250 jämförelserna ledde till "tolkbara" resultat och dels att inget av dessa 10 resultat visade någon signifikant skillnad mellan "televised and face-to-face instruction".

Det finns alltså all anledning att uppmärksamma validitetsfrågan. Campbell & Stanley (1963, s 175) anger åtta faktorer som kan invalidera ett experiments interna validitet och fyra faktorer, som kan påverka experimentets generaliserbarhet.

Som framgått av experimentbeskrivningen, utnyttjas en faktoriell design. Denna är en komplexare form av Campbell & Stanleys design nr 6 *Posttest-Only Control Group Design* (se s 178). Med denna design kontrolleras alla de åtta felkällorna som skulle kunna invalidera experimentets interna validitet, vilken är ett sine qua non för tolkningen av experimentets resultat.

11.2 Yttre validitet

Eftersom det ej förekommer något förtest existerar inte heller interaktionseffekter mellan test och påverkan. Genom slumpmässigt urval av lkk är också interaktionseffekten mellan selektion och påverkan kontrollerad. Däremot är det tveksamt i vilken utsträckning interaktionseffekten mellan den experimentella anordningen och påverkan är kontrollerad inom ramen för denna design. Diskussionen i detta kapitel kommer därför att huvudsakligen beröra interaktionseffekten mellan den experimentella anordningen och påverkan. Denna felkälla kallas också *Hawthorne-effekten* (så benämnd efter de arbetsplatsstudier som Mayo m fl vid slutet av 1920-talet utförde på Hawthorne-företaget; se Hofstätter, 1957, s 37). Den karakteriseras av möjligheten att fp beteende delvis kan vara påverkat av *kunskap* om experimentet och på vilket sätt fp förväntas respondera mot de experimentella stimuli. Med andra ord: *kunskap* om deltagandet i ett experiment kan framkalla beteenden som ej skulle uppstå i vanliga fall.

Individens *kunskap* om deltagandet i experimentet kan naturligtvis ha medverkat till ändring av lk reaktion genom den experimentella påverkan. I detta fall kan experimentets resultat inte uteslutande hänföras till påverkningseffekten.

Invaliderande för experimentets externa validitet kan dessutom de krav vara som experimentsituationen ställer på fpp, omfattande alla tecken som kan förmedla experimentets målsättning (de hypoteser som ställts) till fpp och på så sätt influera individens beteenden. I vilken utsträckning den experimentella situationen påverkar fpp svarsreaktioner, är alltså beroende av graden av bekantskap med experimentets målsättning hos fpp.

Ytterligare förklaringar till fpp förändrade svarsreaktioner skulle kunna vara:

1. fpp är motiverade/ej motiverade på grund av sin värdesättning av vetenskapliga mål och experimentella undersökningar
2. fpp är motiverade på grund av att de önskar visa *socialt önskat beteende*, dvs de önskar göra *rätt* och få bra omdöme om sig.
3. fpp grad av motivation att "*spela sin roll*" under experimentatorns kontroll, dvs i vilken omfattning fpp är villiga att underkasta sig den typiska experimentator-försökspersons-relationen. Har individen, uttagen för deltagande i experimentet, gått med på att delta i experimentet, så medges också implicit av individen i fråga att han kommer att i möjligaste mån göra allt som krävs utan att närmare efterforska målsättningen. I experimentet har fpp medgivande inte varit ett krav för uttagningen.

Utnyttjandet av "*placebo*"-*effekten* (skenbar påverkningseffekt) är utan tvekan en relevant åtgärd för att kontrollera denna felkälla men har ej kunnat

utnyttjas för detta experiment. Hawthorne-effekten har i litteraturen övervärderats och undervärderats. Troligen påverkar den knappast experimentresultaten vid mätning av kunskapsprestationer men det krävs fortsatt forskning för kartläggande av Hawthorne-effekten (se Bracht & Glass, 1968, s 460) såväl med hänsyn till kunskapsprestationer som med hänsyn till mätning av personlighetsvariabler.

I det följande kontrolleras experimentets felkällor och redovisas några resultat i samband med experimentets yttre validitet.

11.3 Kontroll av experimentets felkällor

Syftet med ett experiment är att kunna generalisera resultatet till en individgrupp och till sådana situationer, som ej ingår i experimentet. Även om experimentets yttre validitet, dvs möjligheten till generalisering av experimentets resultat, aldrig fullständigt kan säkerställas, skall här avslutningsvis olika felkällor kontrolleras mot bakgrunden av Bracht & Glass' (1968, ss 437—439) uppräkning av möjliga felkällor:

1. *Populationsvaliditet*

1.1 *För experimentet tillgänglig population kontra målpopulation*

Den för experimentet tillgängliga populationen är alla lkk i andra terminen på LHM, den totala individpopulationen, dvs målpopulationen är alla lkk i M 2 över hela landet. Det beskrivna urvalsförfarandet begränsar generaliseringsmöjligheten till enbart kvinnliga lk med studentexamen i M 2. I vad mån man även skulle kunna våga generalisera resultaten till målpopulationen är beroende av SÖ urvalsförfarande för intagning av lkk på landets olika lärarhögskolor. Men experimentets resultat bör dock i första hand enbart generaliseras till den tillgängliga grupp från vilken fpp valdes ut, dvs kvinnliga lkk med studentexamen, intagna ht 1968 och ht 1969 vid lärarhögskolan i Malmö.

1.2 *Interaktion mellan personlighetsvariabler och påverkan*

Eftersom lkk slumpmässigt har fördelats på de olika experimentella grupperna (obs årtal är kopplade) kan denna felkälla anses vara kontrollerad.

2. *Ekologisk validitet* (avser validitet med hänsyn till organismens relation till omvärlden)

2.1 *Explicit beskrivning av den oberoende variabeln*

Generaliseringsmöjlighet och replicering av de experimentella resultaten förutsätter en uttömmande beskrivning och *fullständig* kunskap om alla aspekter med hänsyn till påverkan och den experimentella situationen. I vad mån den i denna studie presenterade beskrivningen kan anses vara tillräcklig, får man väl kanske först veta i samband med en replicering av experimentet.

2.2 *Interferens genom multipla påverkningar*

När man administrerar två eller flera påverkningar till en och samma individ eller individgrupp, uppstår interferenser. Denna felkälla är under kontroll i experimentet, eftersom varje individgrupp endast utsättes för en enda påverkningskombination.

3. *Hawthorne-effekt*

I vad mån Hawthorne-effekten som felkälla påverkar experimentets resultat, kan ej avgöras. Alla lkk har fått *kunskap* om att de deltar i ett försök. Några entydiga förväntningar på resultat av visst slag kan dock inte gärna ha förelegat.

4. *Nyhets- och avbrottseffekt*

Strävan har varit att schematekniskt planera in experimentet i lkk övriga verksamhet. Men ett visst avbrott har trots allt förekommit. Detta avbrott bedömdes av de deltagande lkk vara av mindre betydelse och av lärarna utan störande inverkan, medan de ej deltagande lkk visade en mera kritisk inställning till sådana avbrott. Nyhetseffekten kan anses vara kontrollerad i någon mån genom den faktoriella uppläggningen, men det är inte helt omöjligt, att vissa samspelseffekter är kopplade med nyhetseffekten (jfr medelvärden för formulär F V, frågorna 6, 8 och 10, nedan).

5. *Experimentledareffekt*

I viss mån kan naturligtvis också experimentledarens beteende omedvetet ha påverkat lkk svarsreaktioner. Experimentledarens sätt att administrera experimentet och hans förväntningar kan ha bidragit till unika beteendereaktioner hos lkk. Att experimentet genomfördes i två faser gör denna felkälla mycket trolig, då andra fasen, jämfört med den första, genomfördes mera rutinerat. (Det kan nämnas, att ljudkvaliteten av lkk under experimentets första fas bedömdes mera positivt, sannolikt på grund av att ett större antal tekniska störningar medförde en högre toleransnivå för avvikelser, medan färre störningar under andra fasen troligen också hade en lägre toleransnivå till följd.)

6. *Förtest-sensitivitet*

Eftersom det ej förekom något förtest kan denna felkälla uteslutas. I samband med introspektion och randomiserade grupper behövs inget förtest.

7. *Eftertest-sensitivitet*

Eftertest kan medföra att den experimentella effekten enbart har uppstått på grund av eftertesten. Problemet i samband med eftertest-sensitiviteten är att påverkningseffekten kan vara latent eller ofullständig och enbart uppstå i samband med en experimentell situation och ett administrerat eftertest, medan i en vanlig situation denna effekt aldrig uppstår. Genom den faktoriella uppläggningen av experimentet kan denna felkälla dock anses vara under experimentell kontroll (jfr Campbell & Stanley, 1963, s 178).

8. *Interaktion mellan påverkan och tid*

En mätning av den beroende variabeln vid två olika tillfällen kan ge olika resultat. Att kontrollera denna felkälla kräver en experimentell design som innehåller flera olika mätningstillfällen. En kontroll av interaktionen mellan påverkan och tid kan delvis åstadkommas inom ramen för experimentet, dvs med hänsyn till de olika bedömningstillfällena för grupperna 1 och 3.

11.4 Några resultat i samband med experimentets yttre validitet

Frågan om experimentets externa validitet, dvs i vilken utsträckning resultaten kan generaliseras till (1) *individpopulationen* och till (2) *omgivningen* (dvs i vilken utsträckning samma resultat kan förväntas i andra situationer) kan, som redan nämnts, aldrig besvaras fullständigt.

I vilken utsträckning denna försöksverksamhet har värderats som ett väsentligt moment i lärarutbildningen eller upplevts som ett störande inslag i den allmänna verksamheten belyser i någon mån experimentets ekologiska validitet. Ett attitydformulär utformades och utdelades för detta syfte till: (1) Lärarhögskolans utbildningspersonal, (2) de lkk i M 2 som ej deltagit i experimentet och (3) de lkk som deltagit i experimentet.

11.4.1 Lärarhögskolans utbildningspersonal (Formulär F VII)

Som framgått har experimentet genomförts i två faser. Formulär F VII distribuerades såväl efter första som efter andra fasen till lärarhögskolans personal. Formuläret besvarades därför av samma individer två gånger. För bearbetning av formuläret delades personalen in i olika befattningsgrupper och svarsreaktionerna på formuläret grupperades i olika svarskategorier enligt tabell 14.

Tabell 14. Befattningshavarnas svarsreaktion på formulär F VII

Befattningsgrupp	berörd 69	berörd 70	ej berörd 69	ej berörd 70	ej svarat 69	ej svarat 70	s:a utsända formul. 69	s:a utsända formul. 70
1. Rektor, personal vid ped.-psyk. inst.	10	9	–	7	8	5	18	21
2. Utbildn.-, fortbildn.- och biträd. utbildn.ledare	4	3	2	3	1	–	7	6
3. Lektorer: Högstadium	10	8	32	31	29	12	71	51
4. Lektorer: Mellanstadium	7	9	12	18	16	1	35	28
5. Lektorer: Lågstadium	–	2	8	2	5	1	13	5
6. Lärare vid F&D skolor	–	–	3	–	2	6	5	6
7. Medverkande i experimentet (klassföreståndare m fl)	6	5	–	–	–	1	6	6
8. Övriga	–	1	7	7	1	1	8	9
Summa	37	37	64	68	62	27	163	132

Som framgår ur tabellen är det år 1969 20 % och år 1970 28 % som har varit berörda eller har känt sig berörda av experimentet. Gruppen som svarade *ej berörd* sände tillbaka formuläret, rubricerat "ej berörd". Någon fråga har denna grupp ej besvarat, medan det från gruppen, rubricerad med *ej svarat,* överhuvudtaget inte inkommit något formulär alls trots en skriftlig påminnelse som sändes ut två veckor efter att formulären hade distribuerats. Mot bakgrund av bortfallet år 1969 38 % och år 1970 20 % måste befattningshavares reaktion på genomförandet av experimentet tolkas med en viss försiktighet. Vågar man med stöd av detta resultat påstå att experimentet varit obesvärande för majoriteten av lärarhögskolans befattningshavare? För gruppen som besvarade formuläret ges i tabell 15 aritmetiskt medelvärde (m), standardavvikelse (s) och antal individer (n) separat för åren 1969 och 1970.

De i tabell 15 redovisade frågorna har mätts med en femgradig skattningsskala. Frågorna 8 och 9 är däremot dikotoma och redovisas nedan i form av råpoäng. Höga medelvärden betyder antingen en positiv inställning till experimentet, eller att försöksverksamheten inte upplevdes som störande. Som framgår ur tabellen värderas experimentets betydelse för lärarutbildningen (fråga 1) och lkk deltagande i experimentet (fråga 5) samt betydelsen av en fortsatt ITV-försöksverksamhet för lärarutbildningen (fråga 10) mycket positivt. Experimentet syns f ö ha upplevts som något störande för den övriga verksamheten.

I vad mån experimentet har inverkat på möjligheterna att använda ITV-studion för den undervisande personalen framgår av följande frekvensfördelning (tab 16).

Att experimentet har varit störande ansåg en relativt stor grupp. Genom en

Tabell 15. Skattningsformulär F VII: Forskningsvärdering (personal). Medelvärde, spridning och n-tal

Fråga	Innebörd	1969		1970	
1	Experimentets betydelse för lärarutbildningen	m	4.33	m	4.28
		s	1.57	s	1.51
		n	31	n	29
2	Experimentets inverkan på undervisningen	m	2.70	m	2.71
		s	1.03	s	1.08
		n	27	n	25
3	Experimentets inverkan på undervisningsplaneringen	m	2.83	m	2.85
		s	1.12	s	1.28
		n	28	n	26
4	Experimentets inverkan på arbetsbelastningen	m	2.79	m	2.58
		s	.96	s	1.05
		n	29	n	26
5	Betydelse av lkk deltagande i experimentet	m	4.34	m	4.32
		s	1.60	s	1.55
		n	29	n	31
6	Experimentets inverkan på eleverna från F&D-skolan	m	2.24	m	2.50
		s	.60	s	.72
		n	17	n	22
7	Experimentets inverkan på materialframställningen	m	2.56	m	2.64
		s	.50	s	.75
		n	18	n	28
10	Betydelse av en fortsatt ITV-försöksverksamhet för lärarutbildningen	m	4.00	m	4.21
		s	1.43	s	1.53
		n	28	n	33

Tabell 16. Skattningsformulär F VII. Svarsfördelning för fråga 8. (Försökets inverkan på användningsmöjligheter av TV-studion i undervisningssammanhang)

År	Svarsalternativ			n	m
	mycket hög grad (1)	ganska hög grad (2)	inte alls (3)		
1969	8	4	12	24	2.35
1970	3	12	13	28	2.71

förbättring av ITV-anläggningen har användningsmöjligheterna för år 1970 varit bättre. Skillnaden testades med McNemar-testet (1 och 2 sammanslagna) eftersom befattningshavarna måste betraktas som ett relaterat sample (se Siegel, 1956, ss 63—67). Skillnaden testades som riktad hypotes på 5 %-nivån. Men denna förbättring har uppenbarligen ej haft någon effekt, då det observerande värdet $\chi^2 = .16$) ej är signifikant.

Trots att experimentet schematekniskt integrerades i lkk vanliga undervisning, förekom det, att lkk i vissa fall fick lämna undervisningen för att delta i experimentet. I vad mån lärarna upplevde detta som störande framgår av svarsfördelningen för fråga 9 (F VII) enligt tabell 17.

Tabell 17. Skattningsformulär F VII. Svarsfördelning för fråga 9. (Störde lkk genom att lämna pågående undervisning för att delta i ITV-försöket?)

År	Svarsalternativ			n	m
	mycket hög grad (1)	ganska hög grad (2)	inte alls (3)		
1969	2	5	8	15	2.40
1970	2	7	17	26	2.58

Skillnaden i utbildningspersonalens upplevelse av lkk frånvaro från undervisningen testades för år 1969 och 1970 som oriktad hypotes på 5 % nivån, men det observerade värdet ($\chi^2 = 2.57$) är ej signifikant. Frånvaron upplevdes alltså vid båda tillfällena på samma sätt.

Sammanfattningsvis kan sägas, att vi med hänsyn till svaren på formulär F VII ej lyckades planera försöksverksamheten schematekniskt på ett sådant sätt, att ingen upplevde denna verksamhet som på något sätt störande med hänsyn till de egna uppgifterna. I relation till den totala gruppen (jfr tab 14) tycks experimentets genomförande dock i stort sett ha varit utan störande inverkan på lärarhögskolans övriga verksamhetsformer, dvs experimentet upplevdes ej som en exceptionell företeelse. Detta experiment och lkk deltagande samt en fortsatt ITV-försöksverksamhet betraktas av den del av personalen som besvarade formulär F VII som betydelsefullt för lärarutbildningen.

11.4.2 De lärarkandidater i M 2 som ej deltagit i experimentet (Formulär F VI)

Syftet är att få någon uppfattning om inställningen hos de lkk i M 2 (år 1969 n = 66 och år 1970 n = 68), som ej deltagit i experimentet: (1) hur de värderar ITV-forskningsverksamhet i samband med lärarutbildningen, (2) i vad mån de är intresserade att delta i kommande ITV-försök och (3) i vilken utsträckning de anser att experimentet var störande för utbildningen, dvs i vad mån de anser att fpp har missat väsentliga undervisningsmoment. Svaren framgår av nedanstående översikt. Det numeriska värdet för svarsalternativen har här satts inom parentes.

Tabell 18. Medelvärde, spridning och n-tal för fråga 1

År	1969	1970
m	2.89	2.55
s	1.27	.96
n	57	65

Fråga 1 lydde:
Är Du intresserad av att deltaga i kommande ITV-försök?
Svarsalternativ
(5) mycket intresserad
(4) något intresserad
(3) varken intresserad eller ointresserad
(2) ganska ointresserad
(1) helt ointresserad

Svaret på denna fråga är uppenbarligen mycket återhållsamt och intresset att medverka i denna försöksverksamhet mycket ringa.

Tabell 19. Medelvärde, spridning och n-tal för fråga 3

År	1969	1970
m	2.83	2.83
s	1.25	1.27
n	57	65

Fråga 3 lydde:
Anser Du att lärarkandidater bör beredas möjligheter att se på egna lektioner via ITV?
Svarsalternativ
(4) mycket ofta
(3) ganska ofta
(2) vid enstaka tillfällen
(1) ej alls

Att få möjlighet att kunna se sina egna lektioner via ITV/VR är av intresse för denna grupp, medan gruppen är ganska ointresserad av medverkan i ITV-försöksverksamheten. Resultatet av dessa två frågor kan tolkas så att en medverkan i ITV-försöksverksamheten tydligen för dessa lkk är en abstrakt aktivitetsform, om vilken man inte riktigt kan föreställa sig vad den skulle kunna innebära för ens egen utbildning och vilken nytta den skulle kunna ha för ens egen praktiska lärarutbildning. Å andra sidan vill man gärna se sina egna lektioner via ITV/VR, troligen eftersom ett sådant arrangemang direkt kan relateras till den dagliga utbildningssituationen, onekligen ett positivt resultat för ITV/VR som utbildningsinstrument. Inte alla utbildningsinstrument värderas lika positivt.

Tabell 20. Medelvärde, spridning och n-tal för fråga 6

År	1969	1970
m	3.72	3.72
s	1.75	1.43
n	57	65

Fråga 6 lydde:
Anser Du att lärarhögskolan bör bereda möjligheter till forskning angående ITV användning i lärarutbildningen?
Svarsalternativ
(5) i mycket stor omfattning
(4) i ganska stor omfattning
(3) i varken stor eller liten omfattning
(2) i ganska liten omfattning
(1) i mycket liten omfattning eller inte alls

Denna svarsreaktion är i kongruens med de tidigare redovisade svaren. Den allmänna formuleringen medför att lkk kan svara positivt, dvs att försöksverksamheten bör förekomma i ganska stor utsträckning, så länge det ej finns någon omedelbar "fara" för att man själv måste ställa upp. För att granska hur dessa lkk ställer sig till att eventuellt gå miste om viss undervisning, dvs för att relatera den positiva inställningen (fråga 6) till försöksverksamheten till en konkret situation samt i syfte att granska hur allvarlig kamraternas (fpp) undervisningsförlust bedöms ställdes frågorna 2 och 4.

Fråga 2 lydde: *Skulle Du vilja deltaga i ITV-försöket även om det skulle innebära ökad arbetsbelastning för Dig?* Svarsfördelningen återges i tabell 21:

Tabell 21. Svarsfördelning för fråga 2

År	ja	nej	vet ej
1969	9	34	14
1970	14	38	13

Av de lkk som entydigt tog ställning är det endast en mindre grupp som för närvarande kan tänka sig att acceptera en större arbetsbelastning.

Skulle man kunna få en mera flexibel lärarutbildning, i vilken lkk kan uppleva direkta praktiska konsekvenser av ett deltagande i försöksverksamhet (tillgodoräknande för slutbetyg, poängsättning, utbyte av arbetsuppgifter etc), skulle det kunna tänkas att en mycket större grupp reagerar positivt.

Fråga 4 lydde: *Skulle Du vilja deltaga i försöksverksamheten även om Du tillfälligt skulle behöva lämna undervisningen?* Svarsfördelningen återges i tabell 22:

Tabell 22. Svarsfördelning för fråga 4

År	ja	nej	vet ej
1969	12	32	13
1970	16	40	9

Att eventuellt lämna ordinarie undervisning för att kunna delta i försöksverksamheten avvisas av flertalet i denna grupp. Eftersom det normalt krävs närvaro vid de schemalagda undervisningstimmarna, är lkk reaktion helt i linje med lärarhögskolans uppfattning att deltagande i lektionsundervisningen är avgörande för kvaliteten på den framtida läraren. Parallella aktiviteter får alltså ej förekomma.

Experimentella försöksplaner medför ofta en varierande arbetsbelastning, som i sin tur väcker krav på någon form av kompensation. Dessutom kan man fråga sig om lärarutbildning inte behöver en mycket större individualisering än den som existerar i dag — vem säger att alla lkk behöver lika många timmar på ett visst moment för att kunna bli "*bra*" lärare?

Eftersom schemaläggningen av t ex en experimentell undersökning av denna omfattning nödvändigtvis medför en viss överlappning mellan försöksverksamhet och annan verksamhet, borde det för den framtida pedagogisk-psykologiska forskningsverksamheten vara av intresse att diskutera ett mera flexibelt utbildningssystem i syfte att kunna genomföra lärarhögskolans uppdrag att utbilda och bedriva forskningsverksamhet utan att skapa negativa attityder hos de involverade personerna. För att i viss mån kunna validera fpp uppgifter om i vad mån de ansåg att de gick miste om väsentliga moment i undervisningen, ställdes samma fråga till de lkk som ej deltagit i experimentet.

Fråga 5 lydde: *Anser Du att Dina kamrater som har deltagit i ITV-försöket gick miste om väsentliga moment i undervisningen?* Svarsfördelningen återges i tabell 23:

Tabell 23. Svarsfördelning för fråga 5

År	ja	nej	vet ej
1969	0	27	30
1970	31	8	26

Svarsfördelningen mellan experimentets första fas (1969) och andra fas (1970) skiljer sig avsevärt. Medan fpp, enligt kamraterna, uppenbarligen ej har gått miste om väsentliga moment i undervisningen under experimentets första fas är det vid genomförandet av andra fasen ett omvänt förhållande. Vad som har förorsakat den tvära omkastningen är omöjligt att kartlägga.

Sammanfattningsvis kan man med hänsyn till formulär F VI säga, att de lkk i M 2 som ej deltagit i experimentet såväl 1969 som 1970 är positivt inställda till ITV-försöksverksamhet och önskar se sina egna lektioner via ITV/VR *ganska ofta* — en positiv inställning till mediet i fråga. I sina svar huruvida de eventuellt skulle kunna tänka sig viss frånvaro från undervisning är lkk mycket tveksamma. Denna reaktion är helt i enlighet med rådande praxis på lärarhögskolan. Dessutom är lkk tveksamma inför att eventuellt åtaga sig någon större arbetsbelastning. Risken att de lkk som deltog i experimentets första fas skulle ha gått miste om väsentliga undervisningsmoment bedöms för år 1969 som obefintlig, medan fpp i andra fasen anses ha gått miste om väsentliga moment.

11.4.3 **De lärarkandidater som deltog i experimentet (Formulär F V)**

De lkk som deltog som fpp i experimentet fick i slutet av andra terminen ett liknande formulär (F V), omfattande 10 frågor. Alla frågor (utom fråga 3 a och b) innehåller fem svarsalternativ. Fråga 3 redovisas därför först.

Fråga 3 a lydde: *Ungefär hur många lektioner måste Du gå ifrån på grund av ITV-försöken?* Svarsfördelningen återges i tabell 24 a.

Tabell 24 a. Skattade antalet förlorade lektioner enligt fpp

Grupp	År	Antal lektioner (Variationsvidd)
1	1970	0–20
2	1969	0–05
3	1969	0–06
4	1970	0–08

De schematekniska svårigheterna för planläggningen av experimentets fas II reflekteras tydligt i variationen. År 1969 lyckades integreringen av experimentet i den övriga verksamheten bättre än 1970.

Fråga 3 b lydde: *Tror Du att Du därigenom missade väsentliga moment i undervisningen?* Svarsfördelningen återges i tabell 24 b.

Tabell 24 b. Svarsfördelning för fråga 3 b

År	ja	nej	vet ej
1969	2	42	4
1970	11	20	17

Skillnaden mellan grupperna är, utöver den möjliga orsaken att fpp i andra fasen gick miste om flera timmar, svår att förklara. Genom en jämförelse av detta resultat med resultatet på fråga 5 (F VI) får denna reaktion en viss validering. Det framgår tydligt att såväl fpp 1970 som deras kamrater i M 2 bedömer frånvaron från undervisningen som allvarligare än 1969.

Den experimentella designen och urvalsförfarandet tillåter en utvärdering av F V genom en variansanalys (ANOVA) för att testa, om det finns skillnader med hänsyn till experimentets påverkningsformer. För varje fråga (utom fråga 3 a och b) utfördes en separat ANOVA. Att utföra en ANOVA på experimentets design tillåter en relativt enkel bedömning av påverkningseffekterna i experimentet. Icke signifikanta effekter tillåter att man betraktar denna

grupps (alla fpp) attityder som en enhet, oberoende av påverkningsformen. Resultaten för dessa ANOVA ges i tabell 25.

Signifikanta F-kvoter är som sammanfattande test mycket användbara indikatorer för en systematisk förändring. F-kvoten anger däremot dock ej riktningen i de påvisbara differenserna. Först en noggrann detaljgranskning med hjälp av t ex kontrastanalyser tillåter en meningsfull tolkning av experimentet. Vägledande i granskningen av ANOVA-tabellerna kommer i denna undersökning att bli:

1. *Mönstret i F-testen.* För att undvika att relativt isolerade F-kvoter ligger till grund för resultattolkningen diskuteras enbart *tolkbara mönster.*
2. *Precisionen i F-testen.* Proportionsvarians (PV) som förklaras av de enskilda effekterna anger noggrannheten i mätningen. Som mått på storleken i sambandet mellan de oberoende variablerna och den beroende variabeln anges Hays' $\hat{\omega}^2$*-index* (Hays, 1970, s 407). $\hat{\omega}^2$ är ett proportionsmått, vars numeriska värden varierar i storlek som en funktion av antalet effekter i ANOVA. Som ett mått på effektstorleken (ES) när alla övriga effekter hålles konstanta, anges Cohens "*effect size index*" (f). En liten effekt anges av J. Cohen (1969, s 278) med f = .10. En medelstor effekt motsvarar f = .25 och en stor effekt motsvarar f = .40.
3. *Styrkan i F-testen.* Sannolikheten (g) att en effekt är av en viss storlek determinerar styrkan i respektive F-kvot (Cohen, J., 1969, ss 1—4). Ett lågt g-värde implicerar att den accepterade alternativa hypotesen bör betraktas som ett tvivelaktigt resultat. Ligger ett g-värde omkring .50 för upptäckandet av en viss effektstorlek, är denna effekt föga övertygande i resultattolkningen. *När g ligger omkring .70, blir tolkningen av ES meningsfull.*

Ett signifikant resultat i tabell 25 säger ingenting om *precisionen* ($\hat{\omega}^2$, f) och *styrkan* (g) i F-kvoterna. Det leder enbart till slutsatsen att systematiska effekter existerar, men det behöver inte nödvändigtvis betyda att dessa effekter är av större intresse. Omvänt gäller, att ett starkt statistiskt samband kan existera trots att resultatet (F-testet) icke är signifikant. Graden av statistisk signifikans är nämligen beroende av stickprovets storlek (för diskussion av relationen mellan påverkan, mätresultat och stickprovsstorlek, se Hays, 1970, ss 323—335). Om fpp har samplats och fördelats slumpmässigt på de olika experimentella betingelserna, gäller allmänt, att *styrkan* i en statistisk relation återspeglas i vår förmåga att med kännedom om påverkan kunna minska vår osäkerhet i att predicera individens värden på den beroende variabeln. Är $\omega^2 = 0$ betyder detta, att det inte finns någon möjlighet alls till en prediktion. Är $\omega^2 = 1.00$ betyder detta å andra sidan, att sambandet mellan påverkan och mätvärdet är perfekt (för diskussion se Hays, 1970, ss 381—385, 424, 547—548; Kerlinger, 1967, ss 203—204). Medan Hays' $\hat{\omega}^2$ anger storleksordningen av en

Tabell 25. ANOVA för skattningsformulär FV. Forskningsvärdering (deltagare i experimentet)

Fråga	Variations-orsak	df	MS	F	$\hat{\omega}^2$	f	g
1	T (ITV/VR)	1	3.01	4.80	.04	.16	.35
	H (handledn)	1	.26				
	TH	1	.01				
	I (TH)	92	.63				
2	T	1	1.26				
	H	1	1.26				
	TH	1	3.01	4.53	.04	.22	.33
	I (TH)	92	.66				
4	T	1	3.01				
	H	1	.84				
	TH	1	1.26				
	I (TH)	92	1.31				
5	T	1	2.34	4.44	.03	.15	.30
	H	1	1.76				
	TH	1	2.34	4.44	.03	.21	.29
	I (TH)	92	.53				
6	T	1	.01				
	H	1	.26				
	TH	1	14.26	13.67	.12	.36	.81
	I (TH)	92	1.04				
7	T	1	2.34				
	H	1	.84				
	TH	1	5.51	4.40	.03	.21	.29
	I (TH)	92	1.25				
8	T	1	.42				
	H	1	.38				
	TH	1	15.04	17.70	.15	.43	.86
	I (TH)	92	.85				
9	T	1	.01				
	H	1	1.26				
	TH	1	.94				
	I (TH)	92	.62				
10	T	1	.04				
	H	1	.04				
	TH	1	9.38	13.87	.12	.38	.86
	I (TH)	92	.68				

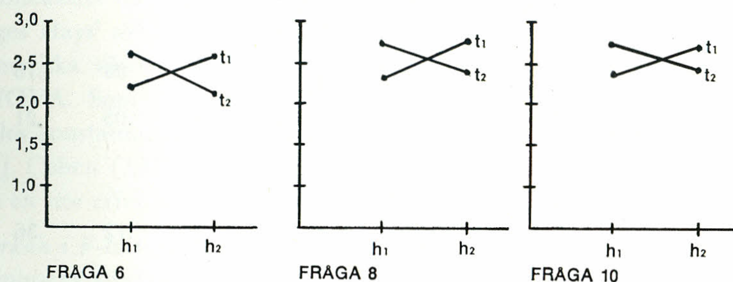
$F_{.99}$ (1, 92) = 7.08
$F_{.95}$ (1, 92) = 4.00

signifikant effekt (PV) i relation till alla övriga effekter, anger g *sannolikheten* att en effekt av viss storleksordning verkligen existerar.

Som jämförelsekriterium har angivits Cohens f. Kvadrerar man f, finner man att f^2 i sig själv kan användas som en approximativ skattning av PV i

individpopulationen (Cohen, J., 1969, s 277). För mera komplexa ANOVA och stora effekter håller denna relation mellan f och g dock ej längre (för diskussion se kap 15.2).

Om den skattade *precisionen* och *styrkan* är relativt liten (bedömningen är subjektiv), kan det därför vara onödigt att satsa tid och arbetskraft på uppföljningen av *triviala* och föga sannolika resultat via t ex efterprövningar enligt Newman—Keuls eller Scheffés förfaringssätt (se Winer, 1962, ss 85—88). En granskning av tabell 25 styrker beslutet att studera enbart sådana resultat som blir signifikanta på $F_{.99}$ (1, 92), eftersom de resultat som är påvisbara på denna signifikansnivå visar en hög precision och en hög styrka. De signifikanta samspelseffekterna på $F_{.99}$ (1, 92), frågorna 6, 8 och 10 åskådliggörs i figur 13.



Figur 13. Grafisk presentation av samspelseffekter (frågorna 6, 8 och 10), där h_1 = handledning, h_2 = ingen handledning, t_1 = självkonfrontation via ITV/VR, t_2 = ingen självkonfrontation via ITV/VR

Som framgår finns det inga skillnader med avseende på huvudeffekterna men däremot föreligger samspelseffekter. Dessa effekter kan vara mycket väsentliga för tolkningen av experimentet. Ett test med ω^2-indexet visar dessutom, som framgår ur tabell 25, att sambandet mellan experimentets oberoende variabler och den beroende variabeln (frågorna 6, 8 och 10) är relativt högt.

Trots att det inom ramen av denna balanserade ANOVA visserligen är möjligt att granska samspelseffekterna, krävs det för en tolkning egentligen en replicering av experimentet. Det bör också anföras, att tolkningen av signifikanta samspelspelseffekter ännu inte är fullständigt kartlagd (Hays, 1970, s 392) samt att ett signifikant F-test inte heller behöver betyda, att man också nödvändigtvis hittar signifikanta värden vid jämförelsen av cellmedelvärdena.

För att testa de signifikanta samspelseffekterna för frågorna 6, 8 och 10 utfördes efterprövningar enligt Newman—Keuls och Scheffés metod (Winer, 1962, ss 77—89). Testet utfördes för $F_{.99}$ (1, 92) = 7.08. Båda testen ledde till samma resultat.

Tabell 26. Kontrastanalys TH, formulär V, fråga 6. (Deltagandet i experimentet var 5 = roligt/1 = tråkigt)

Ordnings-följd	1	2	3	4
Påverkan	t_2h_2	t_1h_1	t_1h_2	t_2h_1
m	3.29	3.42	4.08	4.17
t_2h_2		.13	.79**	.88**
t_1h_1			.66**	.75**
t_1h_2				.09
t_2h_1				
Kritiska värden		.39	.44	.47

Som framgår ur tabell 26 uppger sig alla fpp ha upplevt deltagandet som roligt. Men både de fpp som enbart fick diskutera sina lektioner med en handledare och de fpp som enbart fick se sina lektioner via ITV/VR skiljer sig i sina attityder påvisbart dels från de fpp som ej fick någon påverkan alls, dels från de fpp som utsattes för kombinationen av båda experimentella faktorerna.

Tabell 27. Kontrastanalys TH, formulär V, fråga 8. (Deltagande i liknande försök under lärarutbildningen bedöms som 5 = värdefullt/1 = värdelöst)

Ordnings-följd	1	2	3	4
Påverkan	t_1h_1	t_2h_2	t_2h_1	t_1h_2
m	3.58	3.75	4.42	4.50
t_1h_1		.17	.84**	.92**
t_2h_2			.67**	.75**
t_2h_1				.08
t_1h_2				
Kritiska värden		.35	.40	.43

Som framgår ur tabell 27 indikerar medelvärdena att alla fpp bedömer det som värdefullt att kunna delta i liknande försök under lärarutbildningen. Men såväl de fpp som enbart fick se sina lektioner via ITV/VR som de fpp som enbart fick diskutera lektionen med en handledare skiljer sig från de övriga fpp genom en påvisbart mera positiv attityd.

Som framgår ur tabell 28 anser alla fpp att fortsatt forskning angående ITV/VR-tekniken i stor omfattning bör bedrivas på lärarhögskolan. De fpp som enbart diskuterade med en handledare och de fpp som enbart fick se sina lektioner via ITV/VR skiljer sig från de övriga fpp påvisbart, genom att de indikerar den mest positiva attityden till sådan forskningsverksamhet.

Tabell 28. Kontrastanalys TH, formulär V, fråga 10. (Omfattningen av fortsatt forskning angående ITV i lärarutbildningen borde vara 5 = stor/1 = liten eller inte alls)

Ordnings-följd	1	2	3	4
Påverkan	t_1h_1	t_2h_2	t_1h_2	t_2h_1
m	3.71	3.79	4.38	4.38
t_1h_1		.08	.67**	.67**
t_2h_2			.59**	.59**
t_1h_2				.00
t_2h_1				
Kritiska värden		.31	.36	.38

Utöver de ovan redovisade signifikanta effekterna finns endast små effekter, som, med hänsyn till sambandsstyrkan, ej tycks vara värda en närmare granskning. Eftersom det alltså f ö ej finns någon skillnad mellan de fyra grupperna, kan fpp bedömning av experimentet oberoende av påverkningsformen betraktas som positivt. Medelvärde och spridning för frågorna i formulär F V redovisas i tabell 29.

Tabell 29. Skattningsformulär F V. Forskningsvärdering (deltagare i experimentet). Medelvärde och spridning

Fråga		1 (t_1h_1)	2 (t_1h_2)	3 (t_2h_1)	4 (t_2h_2)
1	m	4.13	4.21	3.75	3.88
	s	1.24	1.28	.94	1.24
2	m	4.21	4.33	4.33	3.75
	s	1.28	1.36	1.35	.94
3		–	–	–	–
4	m	3.33	3.73	3.20	2.79
	s	.87	.99	.75	.80
5	m	4.13	4.17	4.13	3.54
	s	1.24	1.23	1.19	1.17
6	m	3.42	4.08	4.17	3.29
	s	.92	1.21	1.19	.99
7	m	3.67	3.00	2.88	3.17
	s	.96	.72	.72	.70
8	m	3.58	4.50	4.42	3.75
	s	1.01	1.51	1.41	.99
9	m	4.33	4.17	4.42	4.13
	s	1.37	1.34	1.34	1.25
10	m	3.71	4.38	4.38	3.79
	s	1.29	1.41	1.46	.98

Som *sammanfattning* av detta kapitel om experimentets *inre* och *yttre validitet* kan sägas, att den inre validiteten för denna studie genom den ovan presenterade faktoriella designen är säkerställd. Att ett experiment har intern validitet är ett minimikrav utan vilket en experimentell studie över huvud taget inte kan interpreteras. Det anfördes också att generaliserbarheten av experimentets resultat till individpopulationen ifråga är beroende av experimentets externa validitet.

Frågan om den externa validiteten kan dock, liksom frågan om den induktiva inferensen, aldrig besvaras fullständigt. Men det implicita eller explicita syftet med nästan alla experiment är trots allt att också kunna uttala sig om de experimentella resultaten med hänsyn till individer som ej deltagit i experimentet. Experimentets externa validitet granskades med hänsyn till populationsvaliditet och ekologisk validitet, dvs faktorer som definierar individens omgivning och omgivningsrelation. Populationsgeneraliseringen hade varit mycket enkel, om vi hade haft möjlighet att betrakta målpopulationen som tillgänglig population, men det har inte varit fallet. Ur samplingsteoretiska synpunkter bör man därför i första hand enbart generalisera med avseende på experimentets *tillgängliga population.* Men utöver möjligheten att generalisera resultaten med avseende på *målpopulationen* vill man också kunna säga att samma effeker uppnås under andra än de experimentella omgivningsbetingelserna. Denna generalisering baserar sig på antagandet att den experimentella effekten är oberoende av den experimentella omgivningen, dvs man vill kunna säga att experimentet har *ekologisk validitet.*

Utöver diskussionen om experimentets yttre validitet har ovan redovisats några empiriska resultat (formulär F V—VII) som belyser vissa aspekter av experimentets ekologiska validitet, dvs i vad mån experimentet upplevdes som en exceptionell företeelse jämfört med den vanliga utbildningssituationen. De resultat som diskuterats i detta kapitel kan sägas innebära, att experimentet i stort sett inte har upplevts som något speciellt störande inslag i den övriga verksamheten, och att den allmänna attityden överlag är positiv till denna ITV-försöksverksamhet samt till användningen av självkonfrontationstekniken via ITV/VR i lärarutbildningen. Detta resultat talar för mediet och tillåter antagandet att experimentet har ekologisk validitet, eftersom inte enbart de deltagande lkk utan också de ej direkt berörda grupperna reagerade positivt inför denna verksamhet.

DEL 2

ANALYSER AV LÄRARKANDIDATERS SJÄLVBEDÖMNING

12. Problem

Utvecklingen av interna televisionssystem och videobandspelare (ITV/VR-tekniken) tillåter nya forskningsmetodiska tillvägagångssätt för ett systematiskt och kontrollerat studium av interaktionsprocesser. För att studera metodens användningsmöjligheter vid handledning av lkk i undervisningssammanhang utformades den i del 1 beskrivna experimentella undersökningen. En av huvudproblemställningarna som har presenterats i del 1 är: *Vilka effekter har den experimentella påverkan, dvs traditionell handledning och/eller externt förmedlad självkonfrontation via ITV/VR, på lkk självbedömning (perception och värdering)?* Ett välkänt fenomen inom perceptionspsykologin är att individens specifika perceptions- och värderingsstruktur medför individspecifika tolkningar av objekt och händelser. Nedan redovisas en utvärdering av fpp perceptioner och värderingar såsom dessa har blivit manifesterade i skattnings- och värderingsformulär F III.

Ur vetenskaplig synpunkt vet man strängt taget mycket litet om på vilket sätt människor får kunskap om sin omgivning och på vilket sätt perceptions- och värderingsprocesser är relaterade till operationer eller handlingar. Men processen tycks på något sätt sammanhänga med vårt sätt att använda den information vi får förmedlad genom våra sinnesorgan. Problemställningen har behandlats bl a under begrepp som *perceptuell organisation* och *reorganisation av problemmaterial.* Ett grundläggande antagande är att alla processer i samband med förmedlande av information är selektionsprocesser. Hade t ex lkk normstruktur, dvs subjektiva måttstock varit perfekt korrelerad med den aktuella undervisningssituationen, skulle vi å ena sidan via de erhållna ingångsdata ha tillgång till individens normstruktur och å andra sidan kunna beskriva all information som finns tillgänglig i en undervisningssituation.

Vi skulle kunna ange exakt på vilket sätt normstrukturen förändras, när nya parametrar införs eller gamla förändras. Men eftersom fpp äger olika *modeller av verkligheten,* som inte helt sammanfaller med det *objektiva förhållandet,* dvs med mikrolektionen, uppstår informationsförlust eller osäkerhet. Att kunna definiera hela perceptions- och värderingsstrukturen är som nämnts i kap 8, logiskt sett omöjligt, och eftersom vi inte kan specificera perceptions- och värderingsstrukturen fullständigt, vet vi naturligtvis inte heller om vi har

ett mätinstrument av full validitet. Vi måste alltså välja det näst bästa alternativet, nämligen att på ett indirekt sätt skatta instrumentets validitet. Validitetsproblemet diskuteras i kapitel 14.

Varje individ har på varje enskild skala som ingår i formulär F III ett observationsvärde, som mer eller mindre exakt återspeglar individens "sanna" värde. Med vilken precision mätinstrumentet mäter blir föremål för en reliabilitetsgranskning i kapitel 14. Hur "*objektivt*" lkk percipierar och värderar den aktuella undervisningssituationen kommer att behandlas i del 4.

13. Några förutsättningar

Det antas att individen kan värdera och också värderar sig själv som instrument för att uppnå ett mål. Detta antagande medför frågan om vilka kriterier individen använder sig av för bedömning av den egna prestationen. Resultatet av fpp (1) föreställning om i vilken riktning en viss situation förändras när en handlingsplan utförs, (2) perception av situationen i fråga efter att handlingen har utförts samt (3) föreställning om en idealsituation som individen önskar uppnå bestämmer de eventuella förändringarnas storlek och riktning. Resultatanalysen utgör ett försök till skattning av fpp sätt att välja alternativa perceptioner och värderingar och ej av det sätt på vilket fpp väljer alternativa handlingar. För att man skall kunna verifiera en existerande föreställning förutsättes prediktioner. Individen måste alltså testa föreställningens användbarhet (Miller, m fl, 1960, s 175).

På experimentets fenomenologiska utgångsposition baseras således antagandet att det för fpp är mera naturligt att visualisera den tänkta situationen och att fokusera uppmärksamheten på inkongruensen mellan den tänkta situationen och den aktuella situationen än att visualisera ett antal alternativa möjligheter. Man har försökt att på många olika sätt bestämma individens perceptions- och värderingskriterier. En kritisk granskning av de olika "själv-ideal"-kvoterna och kvoterna för självobservation — observation genom andra finns i Wylie (1961, ss 245—249, ss 276—285). Som framgått är den väsentliga problemställningen för denna rapport inte att kartlägga de enskilda individernas perceptions- och värderingsstruktur. Det antas att fpp i de olika grupperna percipierar och värderar (inom slumpmässiga gränser) på samma sätt, dvs innan den experimentella påverkan sattes in, skilde sig grupperna ej påvisbart från varandra. Huvudsyftet med resultatanalysen är att granska om den experimentella påverkan signifikant förändrar individens perceptioner och värderingar. Ett studium av eventuella perceptionsförändringar till följd av de olika påverkningskombinationerna förutsätter att det finns valmöjligheter och att undersökningsplanen innehåller ett tidsintervall. Att fpp skall kunna bete sig ändamålsenligt förutsätter dessutom att olika handlingsplaner eller -strategier existerar. En grundläggande förutsättning för utbildningsprocessen är bl a dels att handlingsplaner, dels att enskilda handlingar eller taktiker framgångsrikt kan förmedlas.

Målsättningen för denna påverkan är att åstadkomma beteendemodifikationer (t ex inlärande av ny färdighet) hos lkk. Denna process påbörjas vanligen med att t ex någon lärare i metodik eller pedagogik förmedlar ett "kommunicerbart" program bestående av pedagogisk-psykologiska och/eller metodiska instruktioner. Lektorn förmedlar i detta syfte mer eller mindre schematiskt (verbalt eller icke verbalt) vad han förväntar sig att lk skall utföra. Men att en viss undervisningsstrategi eller -taktik har framförts i verbal form betyder ej nödvändigtvis att lk har kunnat följa framställningen. Ännu mindre betyder denna kommunikationsform att lk har utvecklat nya perceptions- och värderingskriterier eller "korrekta" undervisningstaktiker, som tillåter genomförandet av en undervisningsstrategi eller -plan. Situationen som skall granskas i resultatanalysen är följande:

Handledaren diskuterar, i en dyadisk konfrontationssituation, med utgångspunkt från en viss undervisningsstrategi ett antal taktiker (konkreta handlingsförslag) i syfte att förmedla sin undervisningsstrategi. Lk utgår däremot från vissa konkreta taktiker (bl a observerade under egen skolgång) och arbetar mot målet att utveckla en strategi. Men även om handledare och lk ofta startar från olika positioner, arbetar de mot samma mål, nämligen att utveckla en lämplig undervisningsplan. Resultatanalysen gäller i första hand att granska om handledarens förmedlande av en viss undervisningstaktik eller strategi har medfört att lkk har utvecklat kriterier, som påvisbart skiljer sig från sådana kriterier som lkk utvecklat på grund av den externt förmedlade konfrontationen med den egna undervisningen.

När det gäller lkk självskattningar kan experimentets frågeställning formuleras om till följande hypoteser:

H_0: traditionell handledning ger ingen effekt
H_0: självkonfrontation via ITV/VR ger ingen effekt
H_0: samspel mellan traditionell handledning och självkonfrontation via ITV/VR föreligger ej.

Vi kan alltså granska i vilken utsträckning den experimentella påverkan har medfört skillnader i lkk färdighet att percipiera och värdera olika undervisningssituationer.

14. Validitet och reliabilitet i mätningen

Den statistiska utvärderingen av experimentet följer den i ruta 6 presenterade analysplanen. Resultatanalysen bygger dels på en a priori-metod, dels på strukturanalysmetoder. A priori-metoder baseras i stort sett på testkonstruktörens uppfattning om ett fenomen, dvs bygger på subjektiva bedömningar och logiska indelningsgrunder. Strukturanalystekniken baseras däremot på matematisk-statistiska modeller som tillåter en statistisk sammanvägning av de observerade variablerna, dvs en empirisk eller operationell definition av dimensionerna i ett mätinstrument. De test som kom till användning är forskningsinstrument och ej standardiserade test. Men tack vare användningen av en kontrollgruppsdesign är en standardisering i detta sammanhang av mindre betydelse (se Cattell & Warburton, 1967, s 6).

När man syftar till att konstruera ett personlighetspsykologiskt test som har validitet och mäter objektivt, krävs det att man redovisar och beskriver instrumentets validitet och reliabilitet. Problem i samband med testets homogenitet och reliabilitet är centrala ämnen i varje psykometrisk lärobok, men enligt Cattell & Warburton (1967, s 6) av sekundär betydelse när det gäller personlighetsforskningen. Författarna menar, att dessa mått är helt ointressanta, om mätinstrumentet är en "inspired creation". Om så är fallet, är dessa mått i allmänhet irrelevanta på grund av att reliabilitetsvariationer vanligen är en funktion av testlängd och stickprovsvarians (se också Lord & Novick, 1968, ss 129—131). Cattell & Warburton (s 6) sammanfattar sin uppfattning på följande sätt:

"... it is more important that a test be well conceived than that it be factor valid; that it be factor valid than item homogeneous; that it be item valid than reliable; and that it be reliable rather than over-standardized on superfluous thousands of cases."

Det är dessutom möjligt att definiera och skatta reliabiliteten på många olika sätt (se Lord & Novick, 1968, ss 198—199), och det har använts olika metoder (ofta utan redogörelse för vilka) för att beskriva ett instruments reliabilitet. Men att redovisa en reliabilitetskoefficient utan ytterligare information är, som framgått, av relativt ringa värde vid en beskrivning av mätinstrumentet ifråga.

I syfte att beskriva eller granska ett mätinstrument börjar man vanligen med en redogörelse för testets reliabilitet och nämner därefter något om instrumentets validitet. Denna tradition, framför allt i den pedagogisk-psykologiska litteraturen, hänför sig kanske till en övervärdering av såväl reliabiliteten som homogeniteten på bekostnad av validiteten. Detta tycks ej vara någon tillfällighet, eftersom man med validitet vanligen avser *innehållsvaliditet* vilken för det mesta ur praktiska synpunkter har bedömts som helt tillfredsställande när det gäller kunskapstest. Överförs denna tradition till den personlighetspsykologiska forskningen, kan det lätt leda till att man föredrar att mäta något som är ointressant men resulterar i mycket höga korrelationskoefficienter i stället för att mäta något väsentligt som medför måttliga reliabilitetskoefficienter. Diskussionen har närmast till syfte att sätta reliabilitetsmåttet i sitt rätta sammanhang (se även Larsson, 1971). Naturligtvis bör man sträva efter att de olika måtten i samband med testbeskrivningen uppnår önskade värden. Även om man i regel kan utgå från den klassiska testteorin, kan en blind tillämpning av modellen medföra felskattningar. Den klassiska reliabilitetsmodellen utvecklades i samband med intelligens- och kunskapstest, där avvikelser från förutsättningarna i regel torde vara små, medan det i personlighetspsykologiska sammanhang kan uppstå betydligt större avvikelser. Eftersom man får det rätta perspektivet på reliabilitetsmåttet först när detta betraktas som en modifierande storhet för validiteten, redovisas först några synpunkter på validiteten i experimentets huvudinstrument F III. (För utförlig diskussion av validitetsproblem hänvisas till Cattell & Warburton, 1967, ss 30—52.)

14.1 Validitet

Formulär F III är huvudinstrument för granskning av i vilken utsträckning de olika påverkningskombinationerna har lett till statistiskt signifikanta effekter. Tillvägagångssättet vid konstruktion och urval av mätinstrument har beskrivits i kap 8.

Om experimentatorn med hjälp av en innehållsanalytisk metod kan fånga in *idiosynkratiska* responser till externt förmedlad självkonfrontation via ITV/VR får han ökade möjligheter att konstruera sådana påståenden som återspeglar de dimensioner som fpp själva väljer ut och som är differentierande i den grad som fpp själva bestämmer.

Den innehållsanalytiska bearbetning, på vilken skattnings- och värderingsinstrumentet F III har baserats, resulterade i en framgångsrik kategorisering och kodning av lkk spontana och simultana kommentarer under återuppspelningen av de videobandade mikrolektionerna. De enskilda kategorierna har sedan formulerats om till de påståenden som utgör mätinstrumentet F III. Man har i olika sammanhang kunnat påvisa att innehållsanalystekniken kan leda till en valid systematisering av verbala kommentarer, sett i relation till individens psykologiska tillstånd. Att empiriskt demonstrera en viss innehållsanalystekniks validitet medför samma svårigheter som validitetsproblemet medför i samband med andra analystekniker (se Gerbner m fl, 1969). Problemet ligger i svårigheten att definiera en viss frågeställning exakt, varför det också är mycket vanskligt att uttala sig om i vilken utsträckning man har åstadkommit en framgångsrik mätning av det som skulle mätas. Med hänsyn till att man i många undersökningar har använt test som ursprungligen konstruerats i annat syfte, bör det poängteras, att mätinstrumentet för denna undersökning har utvecklats på basis av sådana aspekter som de berörda individerna (lkk) anser vara väsentliga i samband med en externt förmedlad självkonfrontation via ITV/VR. Eftersom vi f n ännu ej mera precist kan specificera de enskilda dimensionerna i instrumentet, kan vi naturligtvis ej heller mera direkt, dvs på ett tekniskt sätt, ange om instrumentet framgångsrikt mäter vad vi önskar mäta. Det vi önskar mäta är de 6 olika subjekt-objekt-relationerna som har presenterats i tabell 12. Dessa har beskrivits och definierats i samband med utvecklingen av kategorisystemet. Även om instrumentet med hänsyn till innehållsanalysen enbart kan sägas ha innehållsvaliditet, så måste denna betraktas som en väsentlig och mycket önskvärd validitetsform som, när det gäller personlighetspsykologiska test, ofta saknas eller på mycket lösa grunder påstås finnas. Innehållsvaliditeten bygger på a priori konstruerade dimensioner. Innehållsanalysrapporten, som nämndes ovan, måste betraktas som en solid basis för påståendet att mätinstrumentet har innehållsvaliditet. Att indikera innehållsvaliditeten för instrumentet med hjälp av dimensionsanalytiska metoder ingår i undersökningen. Som en allmän basis för instrumentets validitet betraktas dess användbarhet för en systematisk mätning och en etablering av regelbundna relationer. Mätinstrumentet konstruerades för att mäta perceptionen och värderingen av lkk undervisningsbeteenden genom *bedömare* och ej för att mäta vad eleverna lärde sig under lektionen.

14.2 Reliabilitet

Vill man kunna etablera regelbundna relationer eller predicera ett visst beteende förutsätter detta att testet är ett reliabelt mått. Uppfattningen om vad som bör anses vara ett lämpligt reliabilitetsmått är mycket delad bland testteoretiker. Medan bland andra Cronbach (1960) betraktar *testhomogenitet* som det avgörande kriteriet, anser de testteoretiker, som vidareutvecklat multivariata analyser (Burt, 1964; Guilford, 1954; Horst, 1965 etc), att en hög homogenitet ej är önskvärd, eftersom den minskar den multipla korrelationen med ett kriterium. Som möjliga förluster på grund av homogenitet anges av Cattell & Warburton (1967, s 6):

1. en minskning av räckvidden, eftersom de olika items är så lika varandra att alla smälter samman. Detta medför en förlust i variansen.
2. minskning i möjligheten att överföra eller förändra innehållet när man tillämpar testet på olika kulturella undergrupper, dvs det finns ingen "buffert" för att fånga upp sådana effekter.
3. minskning i möjligheten att testa bestämda personlighetsdimensioner, eftersom homogeniteten successivt medför en sammanblandning av en allmän faktor med en specifik faktor.

På experimentell basis tycks man hittills ej ha kunnat påvisa *"rena"* (högt laddade) items för en bestämd personlighetsfaktor. Det finns många olika tekniska tillvägagångssätt för att beskriva ett begrepp eller ett test operationellt. Åtminstone två indelningsgrunder skall nämnas här. Man kan använda sig av faktoranalysmetoder, som Cattell & Warburton (1967, s 33) betraktar som huvudmetod för att definiera validitet av ett test i den personlighetspsykologiska och motivationspsykologiska forskningen. Den andra gruppen skulle då bestå av olika korrelationsmått (i klassisk mätteknisk mening) och olika logiska operationer (a priori-definitioner).

Ett reliabilitetsmått som föreligger i samband med fpp självskattningar är kommunalitetsskattningen, h_i^2, för de olika i formulär F III ingående items. h_i^2 kan ses som ett approximativt reliabilitetsmått som bygger på en specificering av olika variansdelar med hjälp av faktoranalystekniken. (Relationen redovisas bl a i Bierschenk, 1971 b, ss 39—41.)

Guttman har visat att R_i^2, den multipla korrelationen i kvadrat mellan x_i och övriga x—1 variablerna, är en nedre gräns för H_i^2 (se Rozeboom, 1966, s 261; Lord & Novick, 1968, s 535). Samtidigt gäller att kommunaliteten inte kan överstiga reliabiliteten (R_{ii}). Sammanfattningsvis gäller

$$R_i^2 \leq H_i^2 \leq R_{ii}$$

Både R_{ii} och i synnerhet R_i^2 har använts som skattning av H_i^2, där den senare skattningen blir bättre ju större antalet variabler (x) är, förutsatt att personstickprovet är stort.

Faktoranalyserna har, som framgår ur kap 10.2, utförts på en 3-dimensionell datauppsättning, dvs i analysen har tagits hänsyn till personer, attribut och undervisningstillfällen (testtillfällen). Den dataparallellepiped, som faktoranalyserades, består av n = 96, x = 79 och u = 2. För att kunna analysera dessa tre kategorier med hjälp av en konventionell faktoranalysmodell är det nödvändigt att reducera den 3-dimensionella datauppsättningen till en 2-dimensionell datamatris. Experimentets datauppsättning har lagts ut så, att olika undervisningstillfällen betraktas som *olika, distinkta personer.* (För diskussion se Horst, 1965, ss 315—324.) Som framgått ur kap 10.2 har vi utfört sju olika faktoranalyser, korresponderande till sex olika snitt genom dataparallellepipeden samt en övergripande analys. I bilaga 2, blå, Bierschenk (1972 e), tabellerna 1—6 redovisas h_i^2 (summan av kvadrerade faktorladdningar för item i) för den övergripande analysen. Ligger ett kommunalitetsvärde omkring .30, medför detta att variabelns unika varians (specifik varians + felvarians) är .70. Detta resultat bör man tolka så att variabeln ur reliabilitetssynpunkt är av "ringa" värde, eftersom även r_i^2 i denna analys visar relativt låga värden. Om ej något annat anges, refereras i denna del till de blå bilagorna som har redovisats i Bierschenk, 1972 e. Trots att h_i^2-värden för undervisningstillfälle 1 och 2 också föreligger, redovisas dessa ej, eftersom skattningarna måste betraktas som osäkra och, som en kontroll visade, för höga, på grund av att de är baserade på ett för litet antal fpp i relation till det stora antal items som ingår i testet.

De ANOVA-tabeller som kommer att utvärderas nedan tillåter en mycket detaljerad analys. Med hjälp av faktor V är det möjligt att granska effekterna av den experimentella påverkan på itemnivå. För att resultatredovisningen ej skall bli alltför detaljerad genomförs analysen enbart för de enskilda a priori definierade variabelområdena, dvs *summavariablerna.*

I syfte att få en skattning av reliabiliteten för de enskilda variabelområdena beräknades för varje område och varje undervisningstillfälle Cronbachs "coefficient alpha" (α_C) (Lord & Novick, 1968, ss 87—90). Beräkningen av reliabiliteten genom denna koefficient kräver ej att man först måste bestämma reliabiliteten för ett enskilt item utan koefficienten kan beräknas direkt på en viss given summopoäng. α_C utgör *en nedre gräns* för den klassiska reliabiliteten, men detta förutsätter att mätfelen är okorrelerade (se Larsson, 1971, s 10). Storleksordningen av α_C är en funktion av dels antalet testitems dels homogeniteten i de olika testitems. Denna koefficient är ej någon mätning av itemhomogenitet utan den bör ses som en estimering av reliabiliteten, baserad på

Tabell 30. Reliabilitet för summavariabel: alfa-koefficient (α_C), genomsnittsvärden för kommunalitet ($\bar{h}^2$) och multipel korrelation ($\bar{r}^2$) samt homogenitet (r_s), perception (a_1)

Variabel-område	α_C u_1	u_2	$\bar{h}^2$ $u_1 + u_2$	r_s	$\bar{r}^2$ $u_1 + u_2$
1 Ego-ego	.72	.79	.46	.97	.58
2 Ego-elev	.70	.58	.47	.97	.59
3 Ego-IPO	.33	.35	.50	.95	.63
4 Elev-ego	.57	.64	.49	.94	.60
5 Elev-elev	.00	.31	.61	.90	.70
6 Elev-IPO	.64	.53	.58	.89	.67

u_1: undervisningstillfälle 1
u_2: undervisningstillfälle 2

både den interna konsistensen och antalet items som utgör ett visst test eller deltest (Rozeboom, 1966, s 412). De beräknade α_C för perception och variabelområdena 1—6 redovisas i tabell 30.

Som framgår ur tabell 30 ger reliabilitetsskattningen genom α_C en tillfredsställande skattning för variabelområdena ego-ego (1), ego-elev (2), elev-ego (4) och elev-IPO (6). För variabelområdena ego-IPO (3) och elev-elev (5) avviker denna koefficient markant. En granskning av orsaken till dessa låga värden visade att det ej finns någon nämnvärd itemkovarians för variabelområdena 3 och 5. De enskilda items som beskriver variabelområdena 3 och 5 har tydligen medfört relativt heterogena skattningar. Ett lågt värde för variabelområde 5 borde dessutom kunna förklaras med att det enbart består av 4 items. Anmärkningsvärt är därför det höga värdet för variabelområde 6 trots att detta enbart innehåller 4 items. Variabelområde 6 utmärkes av ett högt värde för itemkovariansen. Både den genomsnittliga multipla korrelationen ($\bar{r}^2$) och den genomsnittliga kommunaliteten ($\bar{h}^2$) redovisas i tabell 30 ovan. Som jämförelsekriterium redovisas också en homogenitetsskattning enligt Spearman—Browns "prophecy"-formel (Rozeboom, 1966, s 408) men med alla reservationer för dess förutsättningar. För beräkningen av formeln har använts $\bar{r}^2$. Denna homogenitetsskattning implicerar att α_C för variabelområdena 3 och 5 ej tycks vara någon lämplig skattning.

I blå bilaga 2 (tabellerna 1—6) redovisas h_i^2-värdena för varje enskilt item samt respektive r_i^2-värden, baserade på båda undervisningstillfällena. Vilka items som med hänsyn till respektive items h_i^2-värde och r_i^2-värde syns vara tillfredsställande anges med (+). De items som ur reliabilitetssynpunkt är tvivelaktiga anges med (—) i tabellens sista kolumn. Avgörandet huruvida ett visst item bör anses ha otillfredsställande gemensam varians baseras på kriteriet $h_i^2 \leq .30$. I blå bilaga 2 redovisas samtidigt värdena för värdering (a_2), medan resultatet för α_C, värdering (a_2) redovisas i tabell 31.

Tabell 31. Reliabilitet för summavariabel: alfa-koefficient (α_C), genomsnittsvärden för kommunalitet ($\bar{h}^2$) och multipel korrelation ($\bar{r}^2$) samt homogenitet (r_s), värdering (a_2)

Variabel-område	α_C u_1	u_2	$\bar{h}^2$ u_1+u_2	r_s	$\bar{r}^2$ u_1+u_2
1 Ego-ego	.00	.12	.45	.96	.56
2 Ego-elev	.16	.12	.45	.97	.57
3 Ego-IPO	.37	.43	.43	.94	.56
4 Elev-ego	.74	.80	.58	.96	.69
5 Elev-elev	.00	.00	.54	.89	.66
6 Elev-IPO	.32	.24	.38	.82	.53

u_1: undervisningstillfälle 1
u_2: undervisningstillfälle 2

Som framgår ur tabell 31 visar α_C för värdering (a_2) med undantag för elev-ego-relationen mycket låga värden. Detta resultat har uppkommit genom att itemkovariansen är nästan lika med eller överstiger summavariansen (α_C blir negativ) vilket implicerar att summavariabeln är oreliabel. Men eftersom det för värderingsaspekten finns ett antal items där fpp skattning är klichéartad, är α_C ett klart olämpligt mått med hänsyn till att den är en nedre gräns för reliabiliteten. Som framgår ur tabell 31 visar nämligen homogenitetsskattningen och den genomsnittliga multipla korrelationen tillfredsställande värden. Resultatet bör tolkas så att fpp värderingar av ett antal items är utan någon större variation. Visserligen innebär detta att sambandet är perfekt, men utan variation kan α_C ej fungera som någon lämplig skattning av reliabiliteten i en summavariabel.

Ego-ego-relationen är i formulär F III operationellt definierad genom 22 items. I samband med ANOVA uteslöts 2 items (8, 9) eftersom dessa har två negativa ändpunkter. I bilaga 4:1—4:6 redovisas alltså 20 items. Vilken pol av de olika bipolära skalorna som bör anses vara den positiva bedömdes av tre av varandra oberoende bedömare. Bedömningen redovisas i Bierschenk, 1972 e, blå bilaga 5. Skalvändningen genomfördes enbart för sådana items, där alla tre bedömarna har varit av samma uppfattning. (Item 4 b har av misstag vänts i fel riktning.) Som framgår ur bilaga 2:1 har enbart ett item (a_2) ej uppfyllt kriteriet.

Ego-elev-relationen är i skattningsformulär F III definierad genom 27 items. Vid den variansanalytiska bearbetningen uteslöts 3 items (2, 11, 14) på grund av att dessa har två negativa poler. De återstående 24 items redovisas i bilaga 4:7—4:13. Som framgår av bilaga 2:2 har endast två items för perception (a_1) och två items för värdering (a_2) ett h_i^2-värde som ej uppfyller $h_i^2 > .30$.

Ego-IPO-relationen är i skattningsformulär F III definierad genom 12 items. I ANOVA ingår samtliga. I bilaga 4:14—4:17 redovisas dessa. Som framgår av bilaga 2:3 har endast ett item för aspekt a_1 resulterat i en obetydlig gemensam varians och en jämförelsevis låg multipel korrelation.

Elev-ego-relationen definieras genom 10 items. Alla dessa ingår i ANOVA. I bilaga 4:18—4:21 redovisas alla 10 items. Av bilaga 2:4 framgår att endast ett item för aspekt a_1 har ett h_i^2-värde som är $\leq .30$, dvs relativt låg gemensam varians.

Elev-elev-relationen definieras genom 4 items, som alla ingår i analysen. I bilaga 4:23—4:24 redovisas dessa 4 items. Som framgår av bilaga 2:5, tabell 5, är värdena genomgående höga. Trots att tre av fyra items visar höga h_i^2-värden, visar α_C ett mycket lågt värde. Detta resultat kan bland annat förklaras med att dessa items visserligen har höga h_i^2-värden med hänsyn till hela testet, men, som anfördes ovan, itemkovariansen är mycket låg inom detta variabelområde.

Elev-IPO-relationen är definierad genom 4 items som alla ingår i ANOVA. De enskilda items redovisas i bilaga 4:26. Inget av dessa items (jfr bilaga 2:5, tabell 6) har ett h_i^2-värde $\leq$ .30. Även dessa items visar relativt höga h_i^2-värden. Dessutom visar α_C åtminstone för perceptionen (a_1) ett högt värde med hänsyn till att det endast gäller fyra items. När det gäller värderingen (a_2) har ett item ett h_i^2-värde $\leq$ .30. Dessutom ligger dessa h_i^2-värden (utom det för item 2) mycket lågt. Item 2 är kanske också något specifikt i detta sammanhang. Detta resultat förklarar kanske bl a den låga α_C-koefficienten. Den påverkas dock också av att det endast gäller fyra items.

För skattningen av reliabiliteten i fpp självbedömningar har flera olika mått (α_C, h_i^2, r_i^2, $\bar{h}^2$, $\bar{r}^2$) använts och diskuterats.

Cronbachs α som en skattning av reliabiliteten i en viss summavariabel visade sig i ett antal fall mindre lämplig. Orsaken är troligen den bristande kovariansen i respektive variabelområde. Resultatet skulle kunna tolkas så, att fpp värdering ger uttryck för mycket heterogena reaktioner i fråga om de items som visar varians. Eftersom det gäller självskattningar, så betyder detta att varje fp värderar varje enskilt item på ett mycket individspecifikt sätt, dvs det finns en tämligen stor inkongruens mellan de enskilda fpp värderingsstruktur.

Om däremot avgörandet i vad mån de enskilda variabelområdena är reliabla baseras på den genomsnittliga kommunaliteten och den genomsnittliga kvadrerade multipla korrelationen, så kan man säga, att summavariablerna har tillfredsställande reliabilitet. Dessutom visade en granskning av reliabiliteten på itemnivån tillfredsställande värden.

15. ANOVA av lärarkandidaters självbedömning

De variansanalytiska principer som gäller för en-, två- eller tre-faktoriella designer kan direkt generaliseras till designer av högre ordning, dvs en-faktordesigner kan ses som en replicering i tvåfaktoriella designer och tvåfaktoriella designer kan ses som en replicering inom ramen för trefaktoriella designer etc. Den i kapitel 5 beskrivna designens precisionsfaktorer (aspekt (A), frågor (V)) har införts för att kontrollera dessa faktorers effekter med hänsyn till de olika nivåerna i självkonfrontation via ITV/VR (faktor T) och traditionell handledning (faktor H) samt mikrolektioner (faktor U). Mätningarna i dessa precisionsfaktorer och i faktor U är repeterade, dvs alla nivåer i respektive faktor administreras till varje enskild fp. Dessa faktorer har helt enkelt till uppgift att öka experimentets precision med hänsyn till jämförelserna avseende faktorerna H och T. En uppställning av ANOVA-tabellerna redovisas i blå bilaga 2:6—2:11. Eftersom alla observationer inom en viss cell har utförts under samma påverkningskombination, har alla effekter som är relaterade till den experimentella påverkan hållits konstanta, varför dessa effekter ej bidrar till inom-cellvariationen. Till felvariansen bidrar skillnader mellan fpp före experimentets början, vissa variationer under experimentets gång eller okontrollerade förändringar i den experimentella tekniken och effekter på grund av unika samspel mellan fpp och påverkan etc.

Den utökade designen medför inte enbart ett ökat antal F-test för en och samma datauppsättning, utan den tillåter också att man separerar all potentiell information i experimentets datauppsättning i distinkta delar, dvs varje enskild del återspeglar endast en speciell aspekt av experimentet.

Modellen antar implicit homogenitet i kovariansmatriserna, som är relaterade till de repeterade mätningarna och som ligger till grund för F-testens validitet. Detta innebär ett mycket restriktivt antagande. Homogenitetsantagandet innebär nämligen att datauppsättningen måste motsvara ett i förväg definierat mönster i dessa kovariansmatriser för att det statistiska testet skall kunna betraktas som exakt. Måste man ifrågasätta homogenitetsantagandet för kovariansmatrisen, bör man tillämpa en konservativ, dvs mera restriktiv prövning. Det kritiska värdet i denna prövning definieras genom att man dividerar frihetsgraderna i nämnaren och täljaren med (df—1), där df = antal frihetsgrader för respektive faktor (se Winer, 1962, s 340). Detta test tillämpa-

des vid prövning av F-kvoterna för faktor V och alla samspel som innehåller denna faktor. Dessutom kan man gardera sig mot påtagliga avvikelser från homogenitetsantagandet, som Nortons studier visar (se Lindquist, 1953, ss 78—86), genom att tillämpa en hög signifikansnivå. Även denna säkerhetsåtgärd har tillämpats. Signifikansnivån bestämdes till $F_{.99}$ (1, 92) = 7.08, vilket innebär att risken för fel av typ I, även för utpräglade avvikelser från normalfördelningen, i realiteten ej överstiger .02 (se Lindquist, 1953, s 81).

Nackdelen med denna mera komplexa design är att den kräver ett stort antal F-test vilket i sin tur ökar sannolikheten för att ett visst antal test resulterar i slumpmässiga signifikanser. Eftersom F-test ej är oberoende av varandra är det mycket svårt att skatta antalet F-test som slumpmässigt har blivit signifikanta. Man bör därför ej lägga alltför stor vikt vid vissa isolerade resultat som kanske tillfälligt blivit signifikanta.

Som signifikansnivå för både signifikanta F-test och för utförda kontrastanalyser fastställdes för alla ANOVA-tabellerna $F_{.99}$ (1, 92) = 7.08. Denna signifikansnivå samt de tidigare diskuterade kraven på redovisning av de experimentella data bestämdes för att motverka tendensen att basera relativt ingående resultatdiskussioner på små skillnader och för att undvika möjligheten till långtgående slutsatser grundade på obefintliga skillnader. Denna tendens påträffar man tyvärr alltför ofta i forskningsrapporter. Som exempel kan nämnas Oliveros (1964) resultatdiskussion som är baserad bl a på att 1 eller 2 items ledde till signifikanta skillnader med α = .05. För övrigt hänvisas till den i kapitel 3 presenterade litteraturredovisningen.

Att ange med vilken precision en viss mätning har utförts måste betraktas som ett grundläggande krav i varje resultatredovisning. Av vilken storlek en viss skillnad bör vara för att den skall kunna anses utgöra en väsentlig information har mycket sällan diskuterats vid rapportering av experimentella resultat i beteendevetenskapliga sammanhang. Finns det ej tillfredsställande erfarenheter inom ett visst forskningsområde eller rörande en speciell problemställning, kan det vara svårt att på objektiva grunder definiera en "stor" effekt för den alternativa hypotesen. Att beskriva en effekt som liten, medelstor eller stor kan därför verka som enbart ett tillägg av *kvalitativa* adjektiv, som ej bidrar till en meningsfull beskrivning av de vunna forskningsresultaten. Men även om det är problematiskt att definiera storheten för en viss effekt, vill jag kraftigt poängtera att man ej får undvika denna uppgift. Att påstå att man absolut inte har möjlighet att veta hur stor en given effekt bör vara betyder att man ej heller har någon logisk basis för att avgöra hur många observationer som är nödvändiga.

Uttrycker man storleken av en viss effekt i form av prediktorvarians och skulle man få ω^2 = .0099 (ca 1 % av den totala variansen i populationen är systematisk varians), så bör man betrakta detta som en liten effekt. Cohen, J. (1969, s 279) skriver:

"The investigator who is inclined to disregard ES criteria for effects this small on the grounds that he would never be seeking to establish such small effects needs to be reminded that he is likely to be thinking in terms of theoretical constructs, which are implicitly measured without error. Any source of irrelevant variance in his measure (psychometric unreliability, dirty test tubs, lack of experimental control, or whatever) will serve to reduce his effects size *as measured,* so that what would be a medium or even large effect if one could use 'true' measures may be attenuated to a small effect in practice . . ."

Som effekt (ES) av medelstorlek anger Cohen, J. (1969, s 280) en prediktorvarians av ω^2 .0588, dvs inte fullt 6 % av den totala variansen i populationen kan betraktas som systematisk varians. Den operationella definitionen av en stor effekt som Cohen föreslagit innebär en proportionell prediktorvarians av ω^2 .1379, dvs en effekt som är något mer än dubbelt så stor som en effekt av medelstorlek.

Signifikanta F-kvoter är som sammanfattande test mycket användbara indikatorer för en systematisk förändring, men först en noggrann detaljgranskning tillåter en meningsfull tolkning av experimentet. Av detta skäl blir bestämningen av precisionen (sambandsstyrkan) i effekterna av avgörande betydelse. Generellt kan sägas att den mest väsentliga delen av dataanalysen börjar när de experimentella data granskas (med precisionsproblemet i minnet) i syfte att studera vad som förorsakade det experimentella utfallet. Med ledning av precisionen och styrkan i sambandet avgörs alltså om det är meningsfullt att granska och kommentera de enkla effekterna i respektive påverkningskombination. Mot bakgrund av denna diskussion genomfördes resultatanalysen i följande tre steg

1. Tolkning av mönstret i F-testen
2. granskning av precisionen och styrkan i F-testen
3. efterprövningar

Eftersom det kan vara svårt att definiera vad som kan anses vara en information av substantiellt värde, har resultatredovisning utförts med tanke på att experiment i beteendevetenskaplig forskning enbart kan utföras i *relativ* isolering. Denna situation överensstämmer ofta ej med de statistiska krav som läroböcker ställer. Skall forskningen leda till sådana resultat som för utvecklingen vidare, krävs naturligtvis all noggrannhet och alla kontroller som är möjliga i ett experiment, men forskningshypoteser bibehålles eller förkastas ej *"in the cavalier way"* som statistiska läroböcker tycks implicera (Hays, 1970, s 459).

15.1 Mönster i F-testen

För varje variabelområde utfördes en separat variansanalys. I tabell 32 nedan redovisas resultatet för alla sex variabelområdena. F-kvoterna granskades i syfte att beskriva tolkbara mönster. I detta sammanhang kan det vara av

intresse att även F-test ($\alpha = .05$) redovisas eftersom dessa kan bidra till att ett mönster framträder tydligare.

Faktor V har lett till signifikanta effekter inom alla sex variabelområdena. Dessa effekter baserar sig på en medelbedömning i varje enskilt item (aspekt a_1 och a_2) och betyder att medelbedömningen i de enskilda items i hög grad skiljer sig från varandra. Ur tolkningssynpunkt är denna information dock av mindre intresse. Av samma skäl tas ej heller någon hänsyn till de signifikanta samspelseffekter som involverar faktor V. En närmare granskning avses alltså ej för faktorn i fråga. Detta anges i tabell 32 genom att alla variationsorsaker som inbegriper faktor V har grupperats i tabellens nedre del.

Faktor A resulterade också i påvisbara effekter för alla sex variabelområdena. Denna faktor tillåter en differentiering av perceptions- och värderingseffekterna. Av detta skäl återfinns denna faktor som ett separeringsmoment i tabellens övre del. De signifikanta effekterna i faktorn innebär således att fpp perception (a_1) och värdering (a_2), oavsett den i experimentella påverkan sedd över alla items, skiljer sig från varandra.

Blir både huvud- och samspelseffekter signifikanta i en variansanalys, kan samspelseffekterna ofta vara av större intresse för resultattolkningen. En analys av de enkla effekterna i den signifikanta påverkningskombinationen tillåter större tolkningsmöjligheter. Tolkningen av mönstret som F-kvoterna bildar i första delen av tabell 32 kan ske dels genom en granskning av de experimentella faktorernas olika samspel (tabellen granskas kolumnvis), dels genom en granskning av varje variationsorsak sedd över alla sex variabelområdena (tabellen granskas radvis).

Som framgår ur tabell 32 accepteras med hänsyn till den övergripande effekten i faktorerna T och H nollhypotesen för alla sex variabelområdena. Detta är ett mycket anmärkningsvärt resultat. Att faktor T ej resulterade i någon signifikant effekt är kanske väntat. Genom att lkk ej ges någon modell för ett ”bra undervisningsbeteende” har de inte heller någon ledning eller någon mera entydigt formulerad norm som skulle kunna leda dem vid perception och värdering av den egna undervisningsprestationen. Som anfördes i den tidigare diskussionen bygger handledarorganisationen i lärarutbildningen bl a på antagandet att lkk i samband med övningsundervisning och auskultation behöver (utöver den egna upplevelsen) någon form av information, förmedlad via handledaren. Förhoppningen är att feedback via en handledare medför någon form av modifiering i perceptions- och värderingsstrukturen som sedan leder till ett önskat lärarbeteende. Med andra ord har handledarens kommentarer till uppgift att ge lk information om i vilken utsträckning en viss given prestation avviker från den för tillfället gällande normen. Påverkan via en handledare har i experimentet ej resulterat i någon signifikant huvudeffekt, *trots* att vi för detta experiment valde en handledare som ansågs (av både lk och lärare) vara mycket skicklig.

Tabell 32. Lärarkandidaternas självbedömning: Sammanfattning av ANOVA-tabellerna

Variationsorsak	Variabelområde 1	2	3	4	5	6
T						
H						
TH		*		*		
U	**	**	**			
UT						
UH						
UTH						
A	**	**	**	**	**	**
AT			*			
AH					**	*
ATH		*		*	*	**
AU	**		*	**		
AUT			**			
AUH						**
AUTH						
V	**	**	**	**	**	**
TV						
HV					*	
THV					*	
UV		*	*			
UTV						
UHV						
UTHV						
AV						
ATV						
AHV						
ATHV						*
AUV			**			
AUTV					**	
AUHV						
AUTHV						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
V: Skattnings- och värderingsformulär F III
** $F_{.99}(1, 92) = 7.08$
* $F_{.95}(1, 92) = 4.00$

Det kan vara svårt att uttala sig om en accepterad nollhypotes utöver konstaterandet att någon systematisk effekt inte har kunnat påvisas. En möjlig orsak till resultatet skulle kunna vara att fpp ännu ej har lyckats bygga upp någon referensram som tillåter transformering av handledarens information, så att den kan modifiera den existerande perceptions- och värderingsstrukturen.

Följer man faktor T i de enskilda variabelområdena, finner man ett signifikant samspel för AT och AUT (3) samt ett signifikant AUT-samspel (5). För faktor H finns ett signifikant AH-samspel (5, 6) och ett signifikant AUH-samspel (6). Möjligen kan detta resultat förklaras med att handledaren framför allt riktade uppmärksamheten på elevernas reaktioner, medan fpp utan att vara medvetna om detta förhållande närmast inriktade sig på att granska hur de lyckades genomföra den tänkta undervisningen ur ämnesmetodiska synpunkter. Denna tolkning överensstämmer med beskrivningen av handledningsprocessen (jfr kap 5).

Resultatet indikerar dels att den externt förmedlade självkonfrontationen via ITV/VR leder till reaktioner som är determinerade av fpp predispositioner och prediktioner, dels att dessa alterneras i riktning mot den information som fpp utsattes för i samband med den traditionella handledningen. Någon övergripande effekt finns dock ej. Ett signifikant TH-samspel finns för variabelområdena 2 och 4, som inom samma områden också har lett till signifikanta ATH-effekter. På så sätt blir det i princip möjligt att granska vad som har föranlett det signifikanta TH-samspelet, perceptionen eller värderingen. Signifikanta ATH-samspel finns för variabelområdena 2, 4, 5 och 6 vilket tyder på att kombinationen av faktorerna T och H kan ha medfört de flesta perceptions- och värderingsförändringarna.

Faktor A representerar aspekterna. Det enhetliga mönstret över alla variabelområdena är ett väntat resultat. Att denna faktor finns med i analysen beror ej på någon önskan att direkt jämföra perception och värdering, utan faktorn har införts som en precisionsfaktor. Faktor A har till uppgift att avskilja perceptionen från värderingen. En övergripande effekt finns också i faktor U, men enbart för variabelområdena 1—3. Resultatet tycks indikera att fpp oavsett påverkan modifierar sina perceptioner och värderingar från undervisningstillfälle till undervisningstillfälle. Anmärkningsvärt är att denna förändring enbart kan konstateras för ego-dimensionen, dvs fpp modifierar sin referensram när det gäller det egna beteendet mot ett objekt, medan någon sådan modifiering ej kan konstateras för de variabelområden där fpp har till uppgift att relatera andra personers agerande mot den egna personen som objekt.

AU-samspelet har resulterat i signifikanta effekter för variabelområdena 1, 3 och 4, vilket innebär att fpp perception och/eller värdering har förändrats från mikrolektion till mikrolektion.

15.2 Precisionen och styrkan i F-testen

Att det finns statistiskt signifikanta F-test och ett mönster i F-kvoterna som tillåter en meningsfull tolkning av de experimentella data är ej något tillräckligt underlag för att man skall kunna bedöma precisionen i dessa data. Det är dessutom mycket svårt att tolka betydelsen av ett experimentellt resultat enbart på basis av den valda signifikansnivån. Alla signifikanta resultat implicerar inte heller på något sätt en och samma associationsgrad mellan de experimentella faktorerna och mätningen. Som redan har diskuterats implicerar ett statistiskt signifikant resultat visserligen att det finns en systematisk varians, dvs något samband mellan oberoende och beroende variabel, men F-kvoten ger inget mått på sambandets storleksordning. Proportionen av varians (PV) som förklaras av de enskilda variationsorsakerna som resulterade i signifikanta F-kvoter redovisas i tabell 33 i form av $\hat{\omega}^2$.

Som framgår ur tabell 33 ligger proportionen systematisk varians, som förklaras av de enskilda effekterna (med undantag av faktor A) mellan .00 och .03. Enligt den definition av små, medelstora och stora effekter som redovisats ovan, har den experimentella påverkan resulterat i små effekter.

I detta sammanhang bör dock uppmärksammas att $\hat{\omega}^2$ är ett proportionsmått vars numeriska värden varierar i storlek som en funktion av antalet effekter i analysen. Om t ex en faktor med repeterad mätning sätts ur spel blir $\hat{\omega}^2$ för de resterande effekterna större än då $\hat{\omega}^2$ beräknas utan denna sammanslagning. De numeriska värdena måste alltså ses i relation till antalet effekter som prövas. $\hat{\omega}$ bör alltså i faktoriella experiment omtolkas som en *partiell* korrelationskoefficient och $\hat{\omega}^2$ som en proportion, ej av den totala variansen utan av vad som återstår av den totala variansen, sedan den del av variansen som hänför sig till de andra faktorerna och samspelen i analysen har uteslutits (partialiserats ut) (Cohen, J., 1969, s 359).

Eftersom den numeriska storleksordningen för $\hat{\omega}^2$ är beroende av antalet variationsorsaker i analysen, krävs det något jämförelsekriterium för att man på objektiv basis skall kunna bestämma om de värden som redovisas i tabell 33 måste betraktas som obetydliga samband eller om de kanske bör ses som ett väsentligt empiriskt resultat. Cohen, J. (1969, ss 266—276) visar relationen mellan index f^2 (”*effect size index*” kvadrerat) och proportionen av varians i individpopulationen med hjälp av olika index som alla utgör en generalisering av den kvadrerade punktbiseriala korrelationen. Som hjälp vid bestämning av styrkan i F-kvoterna ger Cohen (ss 282—347) en serie tabeller som har genererats för enkla ANOVA (enfaktordesign, randomiserade grupper och lika antal individer i varje stickprov). I det enkla fallet kan man använda sig av ω^2 för beräkningen av f. Men för mera komplexa designer håller denna relation ej längre, eftersom ω^2 storlek är beroende av antalet effekter som prövas. Den anger en viss effekts storlek i relation till alla övriga effekter i analysen.

Tabell 33. $\hat{\omega}$-värden för signifikanta F-kvoter: summavariabel

Variations-orsak	Variabelområde 1	2	3	4	5	6
T						
H						
TH		(.02)		(.00)		
U	.03	.01	.01			
UT						
UH						
UTH						
A	.14	.03	.28	.68	.15	.04
AT			(.01)			
AH					.02	(.01)
ATH		(.01)		(.00)	(.01)	.02
AU	.01		(.00)	.00		
AUT			.01		.00	
AUH					.01	
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
(): Anger $F_{.95}$ (1,92) = 4.00

För små värden av f^2 är ω^2 approximativt lika med f. I enkla fall (enkla ANOVA) är ω^2 på motsvarande sätt enbart något mindre än f^2, dvs f kan i sig själv fungera som ett approximativt sambandsmått och f^2 som ett approximativt mått för PV. För stora ES ($f > .40$) avviker dock f och ω i alltför stor utsträckning från varandra för att kunna ge ett approximativt värde, dvs f^2 och ω^2 avviker i ett sådant fall ännu mera från varandra. För att kunna få ett användbart mått som anger ES i komplexa designer, dvs ett mått på effektens storlek, kan det vara till hjälp att ange f.

Beräknas f, så kan tabellerna också användas i samband med faktoriella designer. Användningen av tabellerna kräver dock en viss försiktighet, eftersom df i nämnaren ej längre överensstämmer med de i tabellsamlingen angivna n-värdena. I faktoriella designer har man i regel ett lägre antal df i nämnaren än de på vilka tabellvärdena bygger. Resultatet blir då i huvudeffekterna, att styrkan som avläses i tabellerna anger en överskattning av *sannolikheten* för storleksordningen i respektive effekt.

För de fall, där experimentatorn befarar att tabellvärdena indikerar en överskattning av styrkan i huvudeffekterna, anger Cohen, J. (s 358) en formel för en underskattning av styrkan. Formeln ger en underskattning av sannolikheten som är större än överskattningen. Metoden anger en garanterat lägre

gräns för styrkeskattningen. För samspelseffekter torde däremot styrkan enligt tabellerna stämma ganska väl. Inom vilket intervall styrkan för huvudeffekterna kan komma att ligga exemplifieras i tabell 34 genom en presentation av dels ett korrigerat styrkevärde, dels ett styrkevärde som ej har korrigerats med hjälp av korrektionsformeln.

Tabell 34. Exempel på en korrigerad och en okorrigerad skattning av styrkan i huvudeffekterna. (Faktor A, tabell 7, blå bilaga 2, Bierschenk 1972 e)

Okorrigerad skattning	Korrigerad skattning
$\alpha = .01$	$\alpha = .01$
e = 1	e = 1
n = 192	n' = 47
f = .35	f = .35
g = > .99	g = .78

e: antal nivåer under en faktor (A) minus 1
g: styrka enligt tabellvärdena

Med hjälp av signifikanstestningen anges på den valda signifikansnivån att ett visst fenomen existerar (nollhypotesen förkastas). Precisionen, dvs PV i en population, har skattats med $\hat{\omega}^2$ och framgår ur tabell 33 ovan. Skattningarna säger dock ingenting om styrkan i dessa värden. Att det blivit möjligt att skatta sannolikheten för ett förkastande av nollhypotesen måste ses som ett väsentligt bidrag till inferensproblemet. Med hjälp av styrkeskattning blir det möjligt att man utöver att ange att en effekt av en viss storleksordning (f eller ω^2) existerar, också kan ange hur stor sannolikheten är att denna effekt kan påvisas. Styrkeanalyserna baseras alltså på följande parametrar: (1) kriteriet för styrkan (g), (2) signifikanskriteriet (α), (3) stickprovsstorlek (n) och (4) effektstorlek (ES). Eftersom dessa 4 parametrar är relaterade till varandra, dvs varje enskild parameter är en funktion av de övriga tre, kan styrkeanalyserna utföras på fyra olika sätt. Är nämligen tre parametrar givna så är den fjärde fullständigt bestämd. I styrkeskattningen kan parametrarna manipuleras så att

1. g varierar som en funktion av α, ES och n
2. n varierar som en funktion av ES, α och g
3. ES varierar som en funktion av α, n och g
4. α varierar som en funktion av n, g och ES

Är ett projekt eller experiment avslutat utförs s k post hoc-analyser. I samband med dessa analyser är (1) signifikanskriteriet (α), (2) ES (f) och (3) stickprovsstorlek (n) specificerade. Dessa parametrar kan direkt användas för

bestämning av styrkan i en given F-kvot såsom diskuterats ovan. I de aktuella resultatanalyserna har g beräknats för huvudeffekter (T, H, U, A) samt samspelseffekter av första och andra ordningen. Frihetsgraderna i täljaren är för alla effekter 1, men df i nämnaren varierar. Antalet observationer för huvudeffekterna är 192, för samspel av första ordningen 94 och för samspel av andra ordningen 47. I tabell 35 redovisas g-värden, varvid följande värden är givna: $\alpha = .01$, $e = 1$, $f = .10, .25, .40$ och n som ovan.

Tabell 35. Styrkevärden (g) och olika ES (f) för huvudeffekter (A), samspelseffekter av första ordningen (AB) och andra ordningen (ABC)

f	Effekt A	AB	ABC
.10	.26	.11	.05
.25	.99	.80	.43
.40	>.99	>.99	.90

Vid planeringen av experiment kan man naturligtvis också studera närmare vilka konsekvenser olika α och n har för g. (För diskussion se Cohen, J., 1969, ss 373—397.)

I samband med resultattolkningen kan det dessutom vara av intresse att få veta hur stor en effekt bör vara för att ge ett önskat g-värde. I tabell 36 anges därför några g-värden mot bakgrund av de ovan specificerade parametrarna.

Tabell 36. ES (f) och olika styrkevärden (g) för huvudeffekter (A), samspelseffekter av första ordningen (AB) och andra ordningen (ABC)

g	Effekt A	AB	ABC
.70	.18	.28	.32
.80	.19	.25	.36
.90	.20	.29	.40
.99	.25	.35	.50

Som framgår ur tabellerna 35 och 36 kräver en stor effekt ett mindre antal individer i ett stickprov än små effekter när ett högt g-värde önskas. För små huvudeffekter (f=.10) och ett högt g-värde (t ex g=.80) skulle för detta experiment 2(586) = 1172 fpp varit nödvändiga (se Cohen, J., 1969, s 374), men sådana experiment är knappast möjliga att genomföra. Resulterar en analys i en låg styrka, borde ett sådant negativt resultat, dvs att man ej kunnat förkasta nollhypotesen, betraktas som tvivelaktigt, eftersom nollhypotesen i detta fall är föga meningsfull. Resultatet i ett experiment som t ex visar en styrka

Tabell 37. Effektstorlek och dennas sannolikhet för lärarkandidaternas självbedömningar

Variationsorsak	Variabelområde 1	2	3	4	5	6
1. Effektstorlek (f)						
T						
H						
TH		(.16)		(.21)		
U	.25	.18	.19			
UT						
UH						
UTH						
A	.35	.17	.73	1.13	.36	.23
AT			(.15)			
AH					.19	(.16)
ATH		(.22)		(.21)	(.21)	.34
AU	.21		(.15)	.20		
AUT			.36		.28	.32
AUH						
AUTH						
2. Styrka (g)						
T						
H						
TH		(.35)		(.62)		
U	>.99	.78	.83			
UT						
UH						
UTH						
A	>.99	.72	>.99	>.99	>.99	.95
AT			(.30)			
AH					.52	(.35)
ATH		(.33)		(.30)	(.30)	.76
AU	.62		(.30)	.57		
AUT			.81		.55	
AUH						.69
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
(): Anger $F_{.95}$ (1, 92) = 4.00

omkring .50 för upptäckandet av en viss effektstorlek är föga beviskraftigt eller övertygande när man ej kunnat förkasta nollhypotesen (se Cohen, J., 1969, s 350).

En sammanfattning av effektstorleken (f) för de signifikanta F-testen och de okorrigerade styrkeskattningarna (g) redovisas i tabell 37. Denna skattning bestyrker ytterligare att en detaljerad analys och tolkning av samspelseffekterna är oväsentlig, eftersom styrkan med några undantag (3 samspelseffekter) ligger omkring .50. När g ligger omkring .70, börjar tolkningen bli meningsfull.

15.3 Efterprövningar

En utvärdering av experimentella data med hjälp av F-statistiken betyder att man testar nollhypotesen (H_0), dvs att $m_1 = m_2 = \ldots m_k$ mot en tvåsidig alternativ hypotes (H_1). Ett förkastande av H_0 implicerar, att den aktuella datauppsättningen innehåller systematiska effekter. F-kvoterna anger på ett sammanfattande sätt (omnibustest) att två eller flera medelvärden är påvisbart skilda från varandra. Ett F-test indikerar däremot ej effekternas riktning och det anger inte heller precisionen i mätningen eller styrkan i testet.

Tillvägagångssättet vid utvärderingen av ANOVA-tabellerna har beskrivits ovan. Som framgått betraktas alla statistiska analysförfaranden som bidrar till en reducering av irrelevant variabilitet i de erhållna observationsdata såsom mycket välkomna och informativa hjälpmedel för att öka experimentets styrka.

En resultatredovisning enligt det ovan beskrivna tillvägagångssättet ökar tillförlitligheten i det statistiska resultatutfallet och ger ett säkrare eller beviskraftigare tolkningsunderlag.

Eftersom precisions- och styrkeskattningarna för de signifikanta F-kvoterna i den aktuella analysen implicerar, att det (med några undantag) är föga meningsfullt att ge sig in i en mera ingående tolkning av de enkla effekterna, har de kommentarer som följer i anslutning till kontrastanalyserna (se blå bilaga 3) enbart ett beskrivande syfte. Att kontrastanalyserna trots allt redovisas i form av en bilaga betyder, att vi för fullständighetens skull önskar följa upp de signifikanta F-test, som är baserade på $F_{.99}(1, 92) = 7.08$ (i tabell 32). Ytterligare ett skäl är att bereda den intresserade läsaren tillfälle till en detaljerad granskning av differenserna i medelvärdena för de enskilda samspelseffekterna.

Vid tolkningen av kontrastanalyserna bör man observera att differenserna i cellmedelvärdena granskas oavsett riktningstecknen. För efterprövningar har Newman—Keuls och Scheffés metod använts. Båda metoderna bygger på samplingsfördelningen av cellmedelvärdena (för diskussion se Winer, 1962, ss 80—81). Skulle man vilja dra inferenser från ett sådant förfaringssätt bör man således ta hänsyn till medelvärdenas ordningsföljd.

Med några undantag har efterprövningen ej resulterat i signifikanta differenser. I de fall där det finns påvisbara medelvärdesskillnader mellan perception (a_1) och värdering (a_2) vid skilda lektionstillfällen (u_1, u_2), bör dessa betraktas som en trivial information, eftersom faktor A har införts i analysen i syfte att öka designens precision. Det har ej antagits, att ett sådant resultat skulle innehålla någon meningsfull pedagogisk eller psykologisk information.

Av primärt intresse är sådana medelvärdesdifferenser som hänför sig till t_1 (självkonfrontation via ITV/VR) och t_2 (ingen självkonfrontation) eller h_1 (handledning) och h_2 (ingen handledning) samt dessa effekters olika kombinationer sinsemellan eller med faktor U respektive A. Mot bakgrund av de skäl som angivits ovan är det dock knappast möjligt att formulera hypoteser i samband med post hoc-analyser, dvs beskriva hur ett visst samspel bör se ut.

Resultatet av kontrastanalyserna presenteras avslutningsvis i en sammanfattande tabell (blå bilaga 3, tab 14) både för signifikanta differenser och för enbart tendenser (icke signifikanta skillnader).

Resultatet av analyssteg 2 visar små och medelstora effekter. Bara 3 % eller mindre av den totala variansen i individpopulationen kan betraktas som systematisk varians i de enskilda effekterna. Styrkan i F-kvoterna ligger för samspelseffekterna med tre undantag dessutom på en sådan nivå ($g < .70$) att vi avstår från en mera detaljerad granskning och en mera ingående tolkning av de experimentella data. Det finns tyvärr alltför många exempel på forskningsrapporter där relativt små värden eller relativt isolerade signifikanser gjorts till föremål för ingående tolkningsförsök. Litteraturredovisningen i samband med ITV-forskningen exemplifierar detta faktum. Även om en viss rapportförfattare använder sig av alla försiktighetsåtgärder och reservationer, anges dock implicit riktningen för andras (kanske mindre väl insatta personers) tankeverksamhet; jfr Lees, (1971, ss 3—25) beskrivning av tolkningen av forskningsresultat som framkommit inom ramen för ITV-forskningen.

16. Sammanfattning

Problem och några förutsättningar. Lärarens undervisningsskicklighet och personlighet betraktas a priori som mycket väsentliga faktorer i undervisningssammanhang. Ett välkänt perceptionspsykologiskt fenomen är att individens specifika perceptions- och värderingsstruktur medför individspecifika tolkningar av objekt och händelser. Med en analytisk syn (dock ej i psykoanalytisk mening) på personligheten skulle man kunna anta, att en individs personlighet är uppbyggd av en rad *inlärda* själv-objekt-relationer. Betraktar man dessutom individens *själv* som en organisation av uppfostringserfarenheter (”. . . organization of ways of seeing one's self”; Kerber, 1967, s 24) borde detta innebära att individens olika subjekt-objekt-relationer, t ex synen på sig själv (”self-image”), kan bli föremål för systematiska förändringar. Detta innebär att individen beter sig på ett förutsägbart sätt i olika situationer och vid olika tillfällen. Föremål för resultatredovisning och diskussion i denna del var:

Vilka effekter har den experimentella påverkan, dvs traditionell handledning och/eller externt förmedlad självkonfrontation via ITV/VR, på fpp självbedömning (perception och värdering)?

En variansanalytisk bearbetning kom till användning i denna resultatanalys vid granskningen av om det fanns modifierade perceptions- och värderingsstrukturer till följd av den experimentella påverkan. Men för att kunna granska i vad mån det finns systematiska perceptions- och värderingsförändringar är det nödvändigt att skatta den relativa precisionen i mätningen samt att skatta den relativa säkerheten eller osäkerheten i de funna differenserna.

Validitet. Med mätinstrumentets validitet avses här innehållsvaliditet. Mätinstrumentets variabler baseras på ett kategorisystem som utvecklades i samband med en innehållsanalytisk bearbetning av lkk spontana och simultana kommentarer under externt förmedlade självkonfrontationsprocesser. Att innehållsanalystekniken kan leda till en valid systematisering av verbala kommentarer, sett i relation till individens psykologiska tillstånd, har man kunnat påvisa i olika sammanhang.

Innehållsvaliditet bör betraktas som en väsentlig och mycket önskvärd validitetsform, som ofta saknas i samband med personlighetspsykologiska test.

Eftersom skattnings- och värderingsformulär F III bygger på en innehållsanalytisk bearbetning av lkk simultana och spontana kommentarer under självkonfrontationsprocesser, kan man alltså säga, att formulär F III har innehållsvaliditet i den meningen att det återspeglar sådana dimensioner som lkk själva har utnyttjat.

Reliabilitet är grundläggande för instrumentets validitet. Reliabiliteten kan definieras och skattas på många olika sätt. Två huvudriktningar tycks finnas i dag, (1) de testteoretiker som betraktar homogenitetsmått som det väsentliga reliabilitetsmåttet och (2) de testteoretiker som arbetar med vissa multivariata analystekniker och som betraktar kommunaliteten som mera väsentlig än homogeniteten. Vilken riktning man än ansluter sig till bör man vid bedömningen av mätinstrumentets reliabilitet dock ha följande i minnet: *Den klassiska reliabilitetsmodellen bygger på antagandet att mätfelen från item till item är okorrelerade, en förutsättning som i personlighetspsykologiskt sammanhang kan vara svår att uppfylla.*

För skattningen av reliabiliteten i fpp självbedömningar har flera olika mått (α_C, h_i^2, r_i^2, $\bar{h}^2$, $\bar{r}^2$) använts och diskuterats. Beslutet, att en enskild variabel kan betraktas som reliabel, baserades på h_i^2. De övriga skattningarna tjänade som jämförelsekriterier. Cronbachs alfa-koefficient som en skattning av reliabiliteten i en viss summavariabel (variabelområdena 1—6) medförde i samband med perceptionen (a_1) låga värden för variabelområde ego-IPO (3) och elev-elev (5). När det gäller värdering (a_2) har α_C resulterat i låga värden för variabelområdena ego-ego (1), ego-elev (2) och elev-elev (5). Orsaken till detta är troligen den bristande kovariansen i respektive variabelområde. Den genomsnittliga kommunaliteten och den genomsnittliga kvadrerade multipla korrelationen visar däremot tillfredsställande värden, baserade på kriteriet h_i^2 större än .30. Granskas de enskilda items finner man för perceptionen (a_1) fyra items som ej uppfyller kriteriet, för värdering (a_2) tre items som ligger under gränsvärdena. Sammanfattningsvis kan sägas att reliabiliteten för de enskilda variabelområdena som ANOVA bygger på betraktas som i stort sett tillfredsställande.

En ANOVA utfördes för varje enskilt variabelområde. De aktuella variansanalyserna är av relativt komplex natur. ANOVA-tabellerna upptar tre experiment- och två precisionsfaktorer. De två senare faktorerna fördes in i designen för att öka experimentets precision, dvs reducera felvarians eller osäkerhet som beror på brister i mätinstrumentet. Den utökade designen gjorde det möjligt att granska även mätinstrumentets effekter. Effekterna av faktorerna A och V i samband med faktorerna T, H och U tillåter en separering av den information som hänför sig till mätinstrumentets olika påståenden och aspekter. I resultatanalysen diskuterades dock enbart faktor A och dess olika samspel med faktorerna T, H och U i de fall där de visade sig vara signifikanta med $\alpha = .01$.

Restriktioner och stegvis analys. Som framfördes ovan innebär modellens implicita homogenitetsantagande för kovariationsmatriserna, som är relaterade till de repeterade mätningarna i faktorerna A, V och U, ett mycket restriktivt krav som mycket sällan kan anses vara uppfyllt. För att eliminera riskerna för invalida F-test tillämpades därför några säkerhetsåtgärder som innebär:

1. en konservativ prövning av alla F-kvoter för de variationsorsaker som innehåller faktor V och
2. en hög signifikansnivå, dvs som kritiskt värde bestämdes $F_{.99}$ (1, 92) = 7.08 i syfte att minska risken för fel av typ I.

Det framfördes i diskussionen att signifikanta F-test som visar låga precisions- och styrkevärden ej bör granskas närmare, då ett lågt samband mellan prediktor och kriterium ej tillåter någon prediktion. För att ej riskera att obefintliga eller triviala skillnader görs till föremål för långtgående tolkningar och generaliseringar har ANOVA utförts i tre steg:

1. Tolkning av mönstret i F-testen
2. Granskning av precisionen och styrkan i F-testen
3. Efterprövningar.

Tolkningsresultat. Någon signifikant huvudeffekt i faktorerna T och H finns ej. Sett mot bakgrund av den tidigare förda diskussionen, har informationstillskottet genom antingen T eller H haft samma effekt på fpp perceptions- och värderingsstruktur. Ett anmärkningsvärt resultat! Granskningen av mönstret i F-testen skedde dels med hänsyn till T, H, A eller U-faktorns olika samspel inom ett visst variabelområde dels med hänsyn till en enskild variationsorsak sedd över alla sex variabelområdena.

Faktor T, uppföljd för varje enskilt variabelområde finns i samband med signifikanta samspel i variabelområdena 3 och 5.

Faktor H finns i samband med signifikanta samspel i variabelområdena 5 och 6. Resultatet skulle kunna bero på att handledaren koncentrerade sig på att diskutera *elevernas reaktioner* med fpp. Denna tolkning är kongruent med beskrivningen av handledningsprocessen under experimentet. Fpp *utan* denna påverkan koncentrerade däremot sin uppmärksamhet på *ämnesmetodiska aspekter,* dvs hur de lyckades förverkliga planerade undervisningstaktiker samt den planerade undervisningsstrategin.

TH-samspel blev signifikanta i variabelområdena 2 och 4 vilket också på ett konsekvent sätt resulterade i signifikanta ATH-effekter. Sett över alla variabelområdena finns det i fyra fall signifikanta ATH-samspel (2, 4, 5, 6), vilka är ett tecken på att faktorkombinationen har medfört de flesta systematiska effekterna.

Faktor A visar ett enhetligt mönster med signifikanta effekter i alla variabelområdena.

Faktor U visar däremot signifikanta huvudeffekter enbart i egodimensionen. Detta tyder på att fpp perception och värdering av sig själva (ego-dimension 1—3) i förhållande till omgivningen förändras, medan fpp med hänsyn till omgivningen (elevdimension 4—6) ej kunnat upptäcka någon förändring. Uttryckt på annat sätt betyder detta att den prediktion angående den egna personen, som är baserad på ml_1, testas i ml_2. Testet resulterar i en förändring av fpp självskattning, medan skattningen av elevernas agerande under ml_2 är oförändrad. Denna bild modifieras dock dels genom de redan diskuterade samspelen (där faktor U förekommer) med de experimentella faktorerna och dels genom att samspelet AU inte enbart blev signifikant för variabelområdena 1 och 3 utan också för 4. Sammanfattningsvis kan sägas att faktor U i en eller annan form förekommer i samband med signifikanta effekter i alla sex variabelområdena.

Selektion bland effekterna. Som framgår av ANOVA-tabellerna (blå bilaga 2) skulle 31 olika effekter för vart och ett av variabelområdena ha granskats och tolkats. Alla effekter i analysen var dock ej av lika stort intresse. Ur tolkningssynpunkt har samspelseffekterna mellan experiment- och precisionsfaktorerna varit av större intresse. I de fall, där samspelseffekterna ej tillåtit en separering av fpp perception och värdering med hänsyn till de enskilda beteendeaspekterna, har dessa ej diskuterats. Faktor V har uteslutits dels därför att den är en "medelbedömning", baserad på perception och värdering, en information som blir separat tillgänglig i de samspelskombinationer som innehåller A, dels därför att diskussion ej skulle föras på itemnivå. För att möjliggöra en granskning av fpp perception och värdering av de enskilda items utfördes kontrastanalyser för AV. Dessa har presenterats tillsammans med de enskilda itemgrupperna i Bierschenk, 1972 e, blå bilaga 4, dock utan kommentarer. Att det finns statistiskt signifikanta F-test och ett tolkbart mönster i F-kvoterna kan ej betraktas som en tillräcklig basis för prediktioner.

Precisionen och *styrkan* i mätningen kan ej avläsas ur de signifikanta F-testen. Av detta skäl utfördes analyssteg två, dvs precisionen och styrkan i F-kvoterna granskades. Granskningsresultatet blev (utom i viss mån för faktor A) negativt. De signifikanta F-testen för samspelseffekterna visar en styrka omkring .50. Diskussionen av mönstret i F-testen bör ses mot bakgrund av detta resultat. På grund av utfallet i precisionsgranskningen utfördes ej analyssteg tre i omedelbar anslutning till analyssteg två. Forskningslitteraturen känner alltför många exempel där tolkningen av *små* skillnader blir föremål för långtgående slutsatser. Utan att jag vill göra avkall på de krav som presenterades ovan anser jag dock med hänsyn till experimentets explorativa karaktär, att man ej bör undanhålla sådan information som vid en *försiktig*

tolkning skulle kunna ge viktiga hänvisningar för den fortsatta forskningsverksamheten. Kontrastanalyserna för F-testen ($\alpha = .01$) återfinns därför, försedda med några kommentarer, i blå bilaga 3.

Sammanfattningsvis kan denna resultatanalys knytas an till kap 11.1. Det kan således konstateras, att vi befinner oss i gott sällskap vad beträffar denna resultatanalys. Av de undersökningar med *tillfredsställande* kontrollgruppsdesigner som gjorts inom ITV-forskningen (4 %) visade *inget* undersökningsresultat någon signifikant huvudeffekt. Även om dessa undersökningar gäller *"televised and face-to-face instruction"*, kan man fråga sig om detta är symptomatiskt för ITV-forskningen. Men det återstår att se om samma fenomen också karakteriserar de återstående resultatanalyserna som presenteras i del 3 och 4.

17. Avslutande diskussion

Tillämpas någon form av inferensstatistik vid analysen av beteendevetenskapliga observationsdata förutsätter detta dels (1) att de förutsättningar som en viss statistisk matematisk modell kräver är uppfyllda, dels (2) att man explicit har determinerat styrkan i de statistiska test som har kommit till användning. Det finns knappast någon beteendevetenskaplig rapport utan redovisning av statistiska signifikanstestningar, något som kan synas implicera, att innebörden i sannolikhetsbegreppet skulle vara väl förstått av beteendevetare. Men så är inte fallet. Som redovisades i del 1 finns det tyvärr alltför många exempel på att man varken har försökt att ange, i vilken utsträckning förutsättningarna för en given analysmetod är uppfyllda, eller har försökt att bestämma styrkan i dessa test. Cohen, J. (1969, s 1) uttrycker samma iakttagelse på följande sätt:

"... if we take as evidence the research literature, we find that statistical power is only inferently understood and almost never determined."

Vid utvärderingen av fpp självbedömningar har vi försökt att explicit påvisa, i vilken utsträckning analysmetodernas förutsättningar har uppfyllts, och i vilken utsträckning de statistiska resultatutfallen kan anses vara ett beviskraftigt tolkningsunderlag.

17.1 Huvudeffekter

I samband med den variansanalytiska resultatgranskningen av fpp perception och värdering av de egna mikrolektionerna studerades sex olika kategorier av subjekt-objekt-relationer. När det gäller experimentets faktor T (externt förmedlad självkonfrontation via ITV/VR-tekniken) och faktor H (dyadisk konfrontation under traditionell handledning) har H_0 accepterats för alla sex subjekt-objekt-relationerna.

Någon skillnad i fpp perceptions- och värderingsstruktur som ett resultat av endera faktorn respektive en kombination av den experimentella påverkan har alltså ej kunnat påvisas. Orsakerna till ett sådant resultat (accepterande av H_0) kan naturligtvis vara många, och det är därför svårt att uttala sig om dessa. Skulle man tillåta sig några *spekulativa* antaganden kring dessa resultat, så skulle man kunna förmoda följande:

1. *Självkonfrontation via ITV/VR-tekniken har ej lett till observerbara resultat eftersom*
   (a) påverkningstiden under experimentet har varit för kort för att fpp skall kunna åstadkomma observerbara förändringar
   (b) självkonfrontation ej leder till omedelbara förändringar i fpp perceptions- och värderingsstruktur men däremot till förändringar på längre sikt
   (c) självkonfrontation kräver en systematisk träning i att kunna ta emot och bearbeta *förstahandsinformation,* dvs icke verbalt förmedlad *"själv"*-information
   (d) självkonfrontation medför en temporär desorganisation eller de-automatisering, som i första fasen hos många människor leder till förvåning, rädsla, chock och/eller defensiva positioner inför självkonfrontationsupplevelsen.

2. *Dyadisk konfrontation under traditionell handledning har ej lett till observerbara resultat eftersom*
   (a) handledaren ej har lyckats påverka fpp i högre grad eftersom fpp ännu ej har utvecklat lämpliga testkriterier, dvs pedagogisk-psykologiska normer
   (b) handledare och lk undviker en kritisk analys bl a genom att de använder sig av ord som "a verbal portrait of an individual" (Stoller, 1970, s 11) i syfte att undvika att kritiskt granska det egna beteendet
   (c) lk intar en defensiv ställning mot handledarens kommentarer.

Skulle man följa Salomon & McDonalds (1969, s 14) diskussion, så skulle detta resultat kunna förklaras med att fpp belåtenhet med den egna prestationen *före* handledningens början respektive *före* uppspelningen av de videobandade mikrolektionerna kommer att bestämma vilken information som selegeras, på vilket sätt den värderas och till vilka attitydförändringar värderingarna leder. Men detta är bara spekulationer som kräver en verifiering genom en fortsatt experimentell verksamhet.

Tolkas resultatet rent pragmatiskt eller med utgångspunkt från ekonomiska aspekter, kunde man frestas att rekommendera ett billigare alternativ, dvs lkk behöver under första och andra lärarhögskoleterminen varken handledning och/eller externt förmedlad självkonfrontation via ITV/VR, eftersom den experimentella påverkan ej har lett till någon påvisbar effekt.

Mot en sådan rekommendation står dock dels de möjliga förklaringar som har anförts ovan, dels det faktum att man ej kan dra sådana långtgående slutsatser av ett enda experimentellt utfall. Ett viktigt argument mot ett sådant ställningstagande är dessutom de eventuella långtidseffekterna av den ex-

perimentella påverkan. I samband med experimentets uppföljningsstudier avses att granska närmare i vad mån lkk självbedömning har förändrats, dels sex veckor, dels 24 månader efter experimentets slutförande.

Slutligen visade ett experiment, som har rapporterats av Stoller (1970, s 179), att det är mycket möjligt att beteendeförändringar förekommer (enligt externa bedömare), även när fpp självbedömning såsom den återspeglas i testresultaten ej indikerar någon form av förändring. I vad mån externa bedömare kan konstatera beteendeförändringar kommer att granskas närmare i del 3.

Faktor U representerar i experimentet undervisningstillfällen (ml_1, ml_2). I denna faktor har det kunnat påvisas signifikanta huvudeffekter i ego-dimensionen (variabelområdena 1—3). Detta resultat tyder på att fpp oavsett påverkningskombination har förändrat sin perception och värdering från ml_1 till ml_2. Granskningen av effekternas storlek visade att det i ego-ego-relationen finns en medelstor effekt, medan effekterna i ego-elev-relationen och ego-IPO-relationen är små. Styrkan i dessa effekter implicerar att resultatet ger ett beviskraftigt tolkningsunderlag. Fpp tycks med utgångspunkt från ml_1 predicera sitt eget beteende och testa dessa prediktioner under ml_2. Denna process har i denna analys resulterat i en positivare bedömning av ml_2 än av ml_1.

När det gäller elev-dimensionen kan det ej påvisas någon systematisk förändring. Ett accepterande av H_0 angående elevdimensionen kan ha flera olika orsaker. Att ett predicerat elevbeteende har blivit verifierat under ml_2 kan bero på att de upplevda skillnaderna ej har varit av sådan form och kvantitet att fpp har kunnat upptäcka någon inkongruens mellan bedömningen av elevernas beteende under ml_1 och ml_2. Men det kan också bero på att det saknas mera differentierade kriterier för bedömningen av elevernas beteende eller att fpp i första hand har varit upptagna med den egna personen och således ej haft *"tid"* att studera elevernas beteenden närmare.

Oavsett vad som skulle kunna vara orsaken till detta resultat, borde man undersöka denna problemställning mera ingående, t ex genom ett experiment som omfattar dels en längre tidsperiod dels flera lektionstillfällen än som har varit fallet i detta experiment.

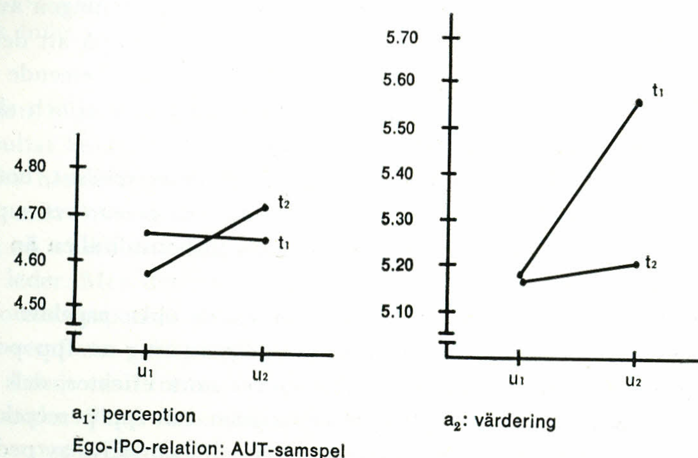
Precisionsfaktor A representerar mätinstrumentets olika aspekter (perception och värdering) och tillåter således en differentiering av fpp perception och värdering. Resultatutfallet visar dels mycket stora effekter, dels en liten effekt i ego-elev-relationen och elev-IPO-relationen. Att fpp perception skulle skilja sig från värderingen är ett väntat resultat och i sig självt av pedagogiskt och psykologiskt föga meningsfullt innehåll. Sannolikheten för de erhållna effekterna är dessutom mycket hög, dock med undantag av ego-elev-relationen, där $g = .72$. Med hänsyn till att skattningen av sannolikheten i huvudeffekterna är en överskattning borde sannolikheten varit högre. Resultatet indikerar att fpp har reagerat positivt på beteendeaspekterna i skattnings-

och värderingsformulär F III och att fpp i den aktuella situationen värderat beteendeaspekterna som väsentliga respektive obesvärande. Ett undantag utgör dock elev-ego-relationen, eftersom fpp värderar "socialt provokativt beteende" (vilket dock ej har förekommit) som ganska besvärande. "Odisciplinerat beteende" hos eleverna (vilket har förekommit) värderas däremot som relativt obesvärande. Skulle man vilja tolka resultatet, så skulle detta kunna betyda, att fpp har en högre toleransnivå när det gäller elevernas agerande sins emellan även om det upplevs som odisciplinerat, medan ett direkt agerande av eleverna mot lk (med ett medvetet eller omedvetet provokativt inslag) upplevs som besvärande.

17.2 Samspelseffekter

De samspelseffekter som har blivit signifikanta med $\alpha = .01$ granskades närmare i Bierschenk, 1972 e, blå bilaga 3. Resultatet har sammanfattats i tabell 14 (bil 3). Precisions- och styrkeskattningar visade, att resultattolkningar endast bör baseras på AUT-samspelet i ego-IPO-relationen, ATH-samspelet i elev-IPO-relationen och AUH-samspelet inom samma variabelområde. Alla övriga effekters styrka implicerar tvivelaktiga resultat.

Vid efterprövningen av AUT-samspelet accepterades H_0, dvs de enkla effekterna skiljer sig ej påvisbart från varandra.



a_1: perception
Ego-IPO-relation: AUT-samspel

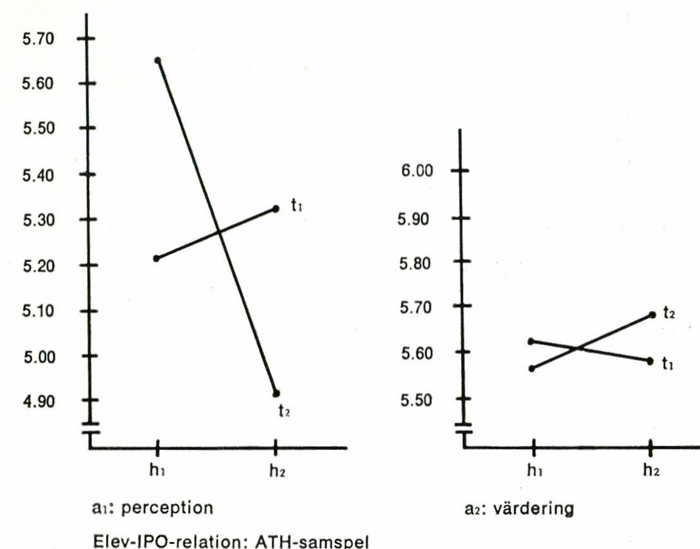
a_2: värdering

Tendensen i medelvärdena angående perceptionen tycks indikera att de fpp som ej fick se sina egna lektioner via ITV/VR (t_2) har förändrat perceptionen från ml_1 till ml_2 i positiv riktning, medan t_1-gruppens perception i det närmaste är oförändrad. I värderingen visar sig tendensen att de fpp, som fick

se sina egna lektioner via ITV/VR (t_1), i samband med ml_2 värderar beteendeaspekterna mera positivt än vid ml_1, medan t_2-gruppens värdering tycks vara relativt oförändrad.

Även om detta resultat måste betraktas som en relativt isolerad effekt, har det dock ett visst intresse för den fortsatta forskningsverksamheten. Det tycks ej vara en orimlig förväntan att i första hand t_1-gruppen borde ha visat en perceptionsförändring, medan t_2-gruppens perception borde ha varit mera konstant. En oväntad tendens är att värderingen har förändrats trots att värderingsstrukturen kan tänkas vara svårare att påverka.

Vid efterprövningen av ATH-samspelet inom elev-IPO-relationen har H_0 förkastats till förmån för H_1. Relationen mellan de enkla effekterna är följande:



a1: perception
Elev-IPO-relation: ATH-samspel

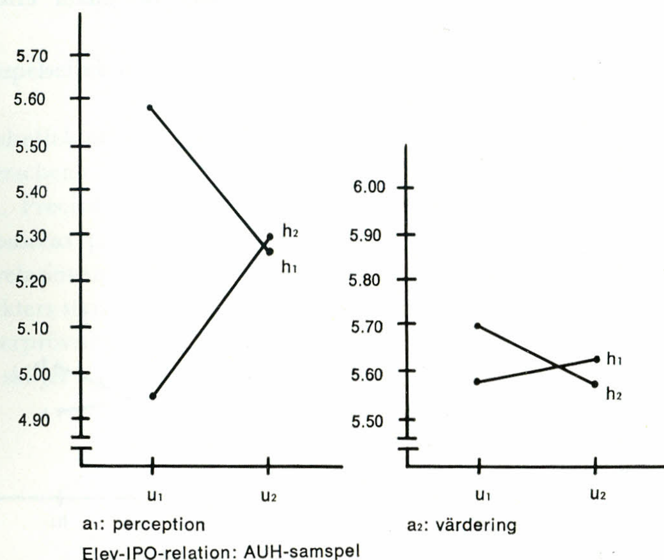
a2: värdering

Den påvisbara skillnaden mellan de enkla effekterna indikerar att de fpp som enbart fick handledning (t_2h_1) i sin perception påvisbart avviker i positiv riktning. Granskningen av precisionen och styrkan i detta resultat visar dessutom att ES är $f=.34$ (medelstor effekt), och att sannolikheten är $g=.76$. Denna effekt kan alltså betraktas som ett relativt säkert resultat.

Värderingen har däremot ej resulterat i någon påvisbar skillnad. Tendensen tycks dock indikera att handledningen har medfört en utjämnande effekt. En intressant tendens återspeglas också i relationen mellan de fpp som ej fick någon påverkan alls (t_2h_2) och de fpp som enbart fick handledning (t_2h_1). Utan någon påverkan tycks fpp visa tendensen att värdera beteendeaspekterna inom elev-IPO-relationen positivare än med någon form av påverkan.

Att handledarens påverkan har resulterat i en signifikant skillnad i perceptionen när det gäller elevernas reaktion på icke-personella-objekt (ämnet, omgivningen etc) kan vara ett resultat av att handledaren har koncentrerat handledningen framför allt kring elevernas beteenden.

Eftersom också AUH-samspelet har resulterat i att H_0 för de enkla effekterna kan förkastas till förmån för H_1, kan vi dessutom granska h_1-effekten närmare. Precisions- och styrkegranskningen visade att ES är medelstor (f=.32) och att g-värdet ligger omkring .70. Relationen mellan de enkla effekterna i AUH-samspelet inom elev-IPO-relationen är följande:



Som framgår ur den grafiska presentationen har skillnaderna mellan h_1- och h_2-gruppen i både perceptionen och värderingen minskat tydligt vid ml_2. Men utjämningen sker genom att förändringen går i olika riktningar. De fpp som ej fick handledning (h_2) är vid ml_1 i sin perception av elevernas reaktioner mot IPO påvisbart mindre positiva än de fpp som fick handledning (h_1). För ml_2 visar sig däremot att $h_1 \approx h_2$.

När det gäller värderingen finns det ej någon påvisbar skillnad mellan de enkla effekterna.

Som framgått av resultatdiskussionen finns det endast tre samspelseffekter, som utgör ett relativt säkert tolkningsunderlag. Signifikanta skillnader mellan samspelens enkla effekter har dock endast kunnat påvisas i ATH- och AUH-samspelen. Men eftersom inga andra effekter tillåter någon tolkning, borde man ej tillmäta dessa relativt isolerade effekter alltför stor vikt.

DEL 3

ANALYSER AV PEDAGOGISKA EXPERTERS BEDÖMNINGAR

18. Bedömningsföremålet "undervisningsprocess"

En granskning och analys av undervisningsprocesser förutsätter observationer. Det antas, att en observation av händelser och/eller objekt till sin natur är subjektiv, dvs enbart ett "subjekt" kan iaktta. Begreppet "subjekt" är dock inte enbart begränsat till intraindividuella processer utan omfattar även subjektets relation till omvärlden, dvs det som är "objektivt".

En viktig målsättning i lärarutbildningen är att utveckla lkk färdighet i att kunnat tolka undervisningsbeteenden och -situationer "realistiskt" eller "objektivt". Att lkk kan tolka en undervisningsprocess förutsätter att det finns en subjektiv "måttstock" i form av ett relationssystem. Individens medvetna perception av sin omgivning och av sig själv i relation till omgivningen baseras alltså på ett relations- eller normsystem. Det antas att denna jämförelse av observationer med ett existerande normsystem bestämmer individens aktuella handlande i form av t ex undervisningstaktiker.

Att ett meningsfullt budskap kan förändra individens föreställningar är en grundläggande premiss för utbildning och utveckling och implicerar samtidigt att individens normsystem är dynamiskt. Att kunna verbalisera och förmedla olika beteendestrategier är en förutsättning för att lärarutbildaren på ett framgångsrikt sätt skall kunna förmedla undervisningsteorier, vilka i sin tur skall bilda ett underlag för en mera objektiv eller realistisk tolkning av undervisningssituationer.

I den ovan redovisade resultatanalysen av lkk självbedömningar (del 2) granskades i vilken utsträckning den inom ramen för en omedelbar undervisningsupplevelse selegerade informationsmängden förändras, på grund av att ny information tillförs från för individen externa källor som handledningskommentarer och/eller via ITV/VR-tekniken. Denna analys ger dock ingen upplysning om i vilken utsträckning de olika experimentella påverkningsformerna har medfört, att lkk tolkar undervisningsbeteenden och -situationer mera realistiskt eller objektivt.

Vill man bestämma objektiviteten i lkk perceptioner och värderingar, är det nödvändigt att man håller undervisningssituationen konstant, så att den kan bedömas på basis av ett yttre kriterium. För den experimentella undersökningen angående den interna televisionens och videobandningens användningsmöjlighet i lärarutbildningen har detta yttre kriterium definierats som en av pedagogiska experter bestående panels medelbedömning. Bedömningspanelen och bedömningsförfarandet har beskrivits i kap 7.2 och 7.3.

I bedömningspanelen ingick 4 pedagogiska experter. Bedömare nr 1 och nr 2, vars observationsdata ligger till grund för medelbedömningen, har bedömt hela protokollmaterialet, medan bedömare nr 3 och nr 4 bedömde icke överlappande delar av materialet. I denna analys utnyttjas därför enbart den datauppsättning, för vilken det finns fullständiga observationsdata. Samstämmigheten i bedömare nr 1 och 2 observationer bedömdes som tillfredsställande (se kap 19.7), varför en analys av bedömare nr 3 och 4 observationsdata i detta redovisningsskede har uteslutits.

I samband med bedömningspanelens skattningar av lkk videobandade mikrolektioner utförde varje bedömare s k principbedömningar. Syftet med sådana skattningar är att få ett värde på experternas mera generella värderingar. En utvärdering av principbedömningar avses dock ej i detta sammanhang.

Att hålla undervisningsprocessen konstant betyder att man måste konservera och återge *exakt* samma situationer och händelser som utgjorde bedömningsmaterialet för lkk perceptioner och värderingar.

Det är allmänt accepterat, att externa observatörer kan åstadkomma större objektivitet vid beskrivning av en persons beteenden. För att man skall kunna studera lkk objektivitet såsom den definieras genom expertbedömningar, krävs det att utgångsmaterialet och skattningsformuläret (kategorierna) är identiska för båda bedömargrupperna.

Med "objektivitet" i tolkningen av undervisningsprocesser avses alltså här en bestämning av i vad mån lkk och pedagogiska experter anger samma numeriska värden vid skattningen av undervisningsprocesser, med hänsyn till de beteendeaspekter som utgör formulär F III. Den operationella definitionen av "objektiva tolkningar" utgörs alltså av bedömaröverensstämmelsen, dels mellan de enskilda pedagogiska experterna, dels mellan experter och lkk.

Ett begrepp som används allt oftare i diskussionen om hur man skall bestämma objektiviteten i olika bedömningar är protokollmaterial. Definitionen av begreppet såsom det skall användas i denna undersökning är följande:

Med begreppet *protokoll* avses en återgivning av beteenden eller "*rådata*" för de beteenden, som skall bli föremål för beskrivningen. I samband med lärarutbildningen kan man urskilja dels *fullständiga* (autentiska) dels *ofullständiga* (selektiva) återgivningar av i utbildningssammanhang relevanta beteenden. Dessa återgivningar kan vara av visuell, auditiv eller skriftlig karaktär. Framställningen av autentiska rådata sker under insamlingsskedet och utgör ett försök att konservera beteenden för att sedan kunna återge dessa så exakt eller tidstroget som möjligt. Ett material av denna typ föreligger i s k *mediaprotokoll*. Denna protokolltyp innefattar inte någon bearbetning eller tolkning. Selektiva rådata genereras antingen i insamlingsskedet genom att olika informationskomprimerande metoder kommer till användning eller genom en efterföljande bearbetning av autentiska data. Selektiva rådata utgörs således

av ett *beteendestickprov*. Detta urval av beteendesekvenser kan ske dels systematiskt, t ex genom olika observationsschemata, dels slumpmässigt, t ex genom subjektiva sammanfattningar. Ett protokollmaterial som utgörs av ofullständiga rådata innefattar ofta någon form av tolkning.

Ett mediaprotokoll ger, jämfört med t ex direktobservationer, bättre underlag för systematiska studier av lkk undervisningsbeteenden. Är det ett riktigt antagande, att lärarens tolkningsfärdighet bygger på en korrekt identifiering av undervisningsbeteende och -händelser, till vilka teoretiska begrepp knyts, så kräver en systematisk utveckling av lkk tolkningsfärdighet, att ett beteende eller en händelse kan göras till föremål för experimentella manipulationer. Mediaprotokoll såsom videoband eller film tillåter såväl en manipulering som en systematisk granskning av undervisningsbeteenden.

I undersökningen användes ett *3-kamerasystem* för att utesluta en redigering av undervisningsprocessen under videobandningen. Inom ramen för de felmarginaler som varje form av audio-visuell upptagning medför, kan alltså undersökningens videobandade mikrolektioner betraktas som identiska med den ursprungliga undervisningen, dvs som ett autentiskt protokollmaterial.

I den resultatanalys som här redovisas granskas reliabiliteten i bedömarnas (1 och 2) perception och värdering av exakt samma protokollmaterial som har varit utgångsmaterial för lkk egna perceptioner och värderingar.

Genom att lkk via ITV/VR kan åstadkomma en extern perception och värdering av sin egen undervisning blir de sina egna "*externa observatörer och kommentatorer*". Denna situation motsvaras av bedömningspanelens situation. Resultatanalys 2 är alltså spegelvänd i förhållande till den första, avseende lkk självbedömningar, varför resultatredovisningen sker enligt samma förfaringssätt. Först redovisas reliabiliteten i bedömarnas skattningar och därefter följer ANOVA. I denna resultatanalys granskas och diskuteras närmare:

1. *Pedagogiska experters samstämmighet vid bedömningen av experimentets protokollmaterial, dels på itemnivå, dels för summavariabler avseende skattnings- och värderingsformulär F III*

2. *Pedagogiska experters bedömning av lärarkandidaternas undervisningsbeteenden som en följd av den experimentella påverkan.*

Hypotesprövningen avser framför allt:
H_0: traditionell handledning ger ingen effekt
H_0: självkonfrontation via ITV/VR ger ingen effekt
H_0: samspel mellan traditionell handledning och självkonfrontation via ITV/VR föreligger ej.

Därtill kommer en diskussion av sådana faktorkombinationer som har resulterat i tolkbara effekter.

19. Granskning av bedömaröverensstämmelse på itemnivå

Människor bedömer människor i en eller annan form. Detta är en vanlig företeelse i lärarutbildningen, trots att bristerna i detta bedömningsförfarande har påvisats (se t ex Sundgren, 1967). Nedan granskas närmare i vilken utsträckning pedagogiska experter överensstämmer i perception och värdering av lkk undervisningsprestationer med hänsyn till de enskilda påståendena i formulär F III. I syfte att erhålla skattningar av överensstämmelsegraden granskades bedömarnas observationsdata med hjälp av en variansanalytisk design. Användningen av ANOVA-designer möjliggör en skattning av sambandet mellan flera olika populationsparametrar. För varje enskilt item i skattnings- och värderingsformulär F III utfördes en ANOVA. Dessutom beräknades Cronbachs "coefficient alpha" (α_C) för varje enskilt variabelområde. α_C redovisas och diskuteras i anslutning till den nedan följande granskningen av korrelationskoefficienterna baserade på ANOVA.

Den grundläggande uppläggningen för beräkningen av olika möjliga reliabilitetskoefficienter är i denna undersökning en tvådimensionell ANOVA-design, som redovisas i tabell 38.

Tabell 38. ANOVA-uppläggning för skattning av olika reliabilitetsmått avseende bedömarnas observationsdata

Index	J	K	U
Antal nivåer	98	2	2
Populationsstorlek	∞	2	2

J: försökspersoner
K: bedömare
U: undervisningstillfällen

Som framgår ur tabell 38 bygger granskningen av reliabiliteten i mätvärdena (från bedömarnas bedömning av lkk undervisningsprestationer) på en fix

ANOVA. Enligt denna ANOVA-tabell skulle man kunna beräkna ett antal olika reliabilitetskoefficienter. En uppställning ges i tabell 39.

Tabell 39. Olika reliabilitetskoefficienter baserade på ANOVA som redovisats i tabell 38

Index	Bedömare (k)	Observationstillfällen (i)	Korrelation r_{ki}
Antal	1	1	r_{11}
	1	2	r_{12}
	2	1	r_{21}
	2	2	r_{22}

Beroende på om man kan anta att populationen innehåller ett infinit antal situationer och bedömare, eller om man (som i experimentet) bör begränsa generaliserbarheten till de aktuella nivåerna, konstrueras testkvantiteterna på olika sätt. I samband med denna undersökning betraktas både bedömarna och lektionstillfällena som fixa. Något slumpmässigt urval har ej varit möjligt. I resultatanalyserna avseende nivåskillnaderna ingår undervisningstillfällena som variationsorsak. Detta innebär, att reliabilitetsgranskningen i första hand gäller överensstämmelsegraden mellan bedömare nr 1 och nr 2 avseende ett lektionstillfälle, dvs r_{21}. r_{22} kan lätt beräknas över Spearman—Browns formel (se Rozeboom, 1966, s 409). För en utförlig diskussion av definitionsproblem och beräkning av olika reliabilitetskoefficienter hänvisas till Fisher, 1946, ss 211—241; Lindquist, 1953, ss 357—372; Winer, 1962, ss 124—128; Medley & Mitzel, 1963, ss 309—316.

Genom en variansanalytisk skattning får vi koefficienter som är skattningar, baserade på två varianser (systematisk varians och felvarians). Skiljer sig den systematiska variansen ej från felvariansen, så är korrelationen ej heller signifikant. Skillnaderna mellan produktmomentkorrelationen och intraklasskorrelationen diskuteras utförligt av Fisher (1946, s 218). För det mesta anges för korrelationskoefficienter enbart punktskattningar. Ett precisionsmått som i samband med reliabilitetsskattningar är av större informativt värde är intervallskattningen. För r_{21} anges därför konfidensintervallen med P ($r_n \leq P_I \leq r_ö$) = .99. (För beräkning, se Fisher, 1946, ss 216—221.) När F-kvoten är liten, blir vanligen värdena för konfidensintervallens nedre gräns negativa. I de fall detta inträffar har r_n fått värdet noll, eftersom P_I inte kan bli negativt. För varje enskilt item inom respektive variabelområde redovisas nedan intraklasskorrelationen och precisionen i detta sambandsmått i form av intervallskattningar. För alla utförda prövningar gäller såsom tidigare α = .01 om ej något annat uttryckligen anges.

19.1 Ego-ego-relation

Tabell 40. Ego-ego-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a_1)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
1	Emotionellt tillstånd	.63	.41	.77
2	Uppträdande	.62	.41	.77
3	Tålamod med eleverna	.38	.11	.60
4	Sinne för humor	.57	.34	.74
5	Röstnyansering	.48	.23	.68
6	Tydlighet i tal	.41	.14	.62
7	Röststyrka	.59	.36	.75
10	Kroppshållning	.36	.08	.59
11	Tics, ryckningar etc	.00	.00	.00
12	Användning av gester	.65	.40	.79
13	Lek med föremål	.61	.40	.77
14	Faktakunskaper	.15	.00	.41
15	Användning av stereotypa uttryck	.57	.34	.74
16	Användning av ofullständiga meningar	.06	.00	.34
17	Användning av språkligt felaktiga uttryck	.28	.00	.52
18	Dialekt i tal	.64	.44	.78
19	Användning av svåra begrepp utan förklaringar	.29	.01	.53
20	Blockeringar (blackouts)	.31	.03	.54
21	Tydlighet i handstilen på tavlan	.35	.07	.58
22	Användning av retoriska frågor	.22	.00	.48

Kritiskt värde ($n = 90, \alpha = .01$) : $r = .267$, Fisher, 1946, s. 209

Som framgår ur tabell 40 visar punktskattningen, att 4 items ej uppfyller gränsvärdet $r_{21} \geq .27$. För ytterligare 4 items (10, 19, 20, 21) kan populationsparametern (kriteriet $r_n < .10$) endast i mycket ringa grad förväntas skilja sig från $r_{21} = .00$. Item nr 21 baseras på $n_1 = 72$, $n_2 = 76$ bedömningar vilket innebär, att ej alla lkk har använt sig av tavlan. För att man skall kunna granska orsaken till ett lågt sambandsmått, redovisas i tabell 41 medelvärde och spridning, dels för de items som i punktskattningen ej uppfyller kriteriet, dels för de items vilkas nedre gräns är $r_n < .10$.

Tabell 41. Ego-ego-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, perception (a_1)

Item nr	Innebörd	ml	m	s	±
10	Kroppshållning	1	4.68	.62	–
		2	4.90	.50	
11	Tics, ryckningar etc	1	7.00	.00	+
		2	6.99	.05	
14	Faktakunskaper	1	4.74	.64	–
		2	5.00	.60	
16	Användning av ofull-	1	5.88	.83	–
	ständiga meningar	2	6.02	.87	
17	Användning av språk-	1	6.38	.71	–
	ligt felaktiga uttryck	2	6.48	.56	
19	Användning av svåra	1	6.44	.21	+
	begrepp utan förklaringar	2	6.43	.19	
20	Blockeringar	1	6.02	.99	–
	(blackouts)	2	6.26	.78	
21	Tydlighet i handstilen	1	4.59	.88	–
	på tavlan	2	4.80	.88	

ml: mikrolektion s : standardavvikelse
m : medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 41 finns det ett antal items som antingen saknar variation eller har mycket låga spridningsvärden. För att man skall kunna separera dessa items från sådana, som visar en godtagbar variation (bedömningen är subjektiv), bestämdes med utgångspunkt från råmaterialet gränsvärdet $s < .50$. Med utgångspunkt från detta kriterium betraktas i denna studie alla items vars s-värden är lägre än .50 som items "utan variation", medan uttrycket "perfekt" överensstämmelse används när $s = .00$.

Att spridningen i bedömarnas skattningar är mindre än i lkk skattningar är en naturlig följd av dels det faktum att bedömarnas skattningar hänför sig till $k = 2$, dels det faktum att bedömarnas referensram troligen är mera utpräglad än lkk. I vad mån de items som redovisas ovan verkligen måste betraktas som oreliabla beror i mycket hög grad på med vilka mått man mäter. Problematiken i reliabilitetsgranskningen skall exemplifieras med hjälp av item nr 11. För detta item är $r_{21} = .00$. Granskas orsaken till detta resultat med hjälp av medelvärden och spridningar, finner man för

$$ml_1 \quad m = 7.00, s = .00 \quad \text{och för } ml_2 \quad m = 6.99, s = .05.$$

Detta resultat indikerar, att bedömarna i samband med ml_1 överensstämmer "perfekt" i skattningen, och att denna skattning för ml_2 mot bakgrund av kriteriet ($s < .50$) inte heller visar någon variation. Sammanfattningsvis kan

man alltså säga, att bedömarna ej har kunnat percipiera tics, ryckningar etc hos lkk (n=96). Att beskriva reliabiliteten, såsom den är definierad, för ett item som ej har någon spridning är alltså ej möjligt, eftersom någon reliabilitet i ett sådant item ej kan beräknas. Men ur tolkningssynpunkt är det ett viktigt resultat, att bedömarna överensstämmer "perfekt" i sin skattning av att ingen av lkk har visat tics, ryckningar etc.

Med denna utgångspunkt granskades tabell 41. Av de 8 items måste 6 betraktas som tvivelaktiga ur reliabilitetssynpunkt, men med hänsyn till den aktuella analysen och med utgångspunkt från punktskattningen visar endast 3 items (tab 40) otillfredsställande värden.

Sammanfattningsvis kan sägas, att bedömarnas perception av de enskilda beteendeaspekterna inom ego-ego-relationen är i stort sett tillfredsställande.

Tabell 42. Ego-ego-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a_2)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
1	Emotionellt tillstånd	.55	.32	.72
2	Uppträdande	.53	.30	.71
3	Tålamod med eleverna	.58	.35	.74
4	Sinne för humor	.34	.06	.57
5	Röstnyansering	.73	.57	.84
6	Tydlighet i tal	.57	.33	.73
7	Röststyrka	.07	.00	.35
.				
.				
10	Kroppshållning	.66	.46	.79
11	Tics, ryckningar etc	.66	.46	.79
12	Användning av gester	.29	.01	.53
13	Lek med föremål	.24	.00	.49
14	Faktakunskaper	.15	.00	.41
15	Användning av stereotypa uttryck	.23	.00	.49
16	Användning av ofullständiga meningar	.18	.00	.45
17	Användning av språkligt felaktiga uttryck	.30	.02	.54
18	Dialekt i tal	.51	.26	.70
19	Användning av svåra begrepp utan förklaringar	.00	–	–
20	Blockeringar (blackouts)	.19	.00	.45
21	Tydlighet i handstilen på tavlan	.32	.04	.62
22	Användning av retoriska frågor	.38	.11	.60

Kritiskt värde ($n = 90, \alpha = .01$): $r = .267$, Fisher, 1946, s. 209

Sju items i tabell 42 uppfyller ej kravet $r_{21} \geq .27$. Därutöver har item nr 4, 12, 17 och 21 en nedre intervallgräns, som innebär, att man kan förvänta sig mycket låga samband. För $_{21} < .27$ och $r_n < .10$ redovisas m och s i tabell 43.

Tabell 43. Ego-ego-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a_2)

Item nr	Innebörd	ml	m	s	±
4	Sinne för humor	1	2.87	.52	–
		2	3.01	.67	
7	Röststyrka	1	5.14	.54	–
		2	5.12	.48	
12	Användning av gester	1	3.53	.77	–
		2	3.30	.76	
13	Lek med föremål	1	5.67	.44	–
		2	5.69	.51	
14	Faktakunskaper	1	5.67	.47	+
		2	5.73	.31	
15	Användning av stereotypa uttryck	1	5.48	.78	–
		2	5.52	.82	
16	Användning av ofullständiga meningar	1	5.67	.43	–
		2	5.62	.55	
17	Användning av språkligt felaktiga uttryck	1	4.53	.65	–
		2	4.51	.55	
19	Användning av svåra begrepp utan förklaringar	1	6.44	.22	+
		2	6.43	.19	
20	Blockeringar (blackouts)	1	4.00	.67	–
		2	4.07	.74	
21	Tydlighet i handstilen på tavlan	1	5.08	.63	–
		2	5.25	.39	

ml: mikrolektion
m : medelvärde
s : standardavvikelse
±: reliabel/oreliabel

När det gäller sambandet i bedömarnas värderingar syns 9 items ha låga reliabilitetsvärden, medan 2 items har så låga spridningar att intraklasskorrelationen måste betraktas som ett olämpligt reliabilitetsmått. Spridningen i item nr 19 ($r_{21} = .00$) uppfyller ej kriteriet. Båda bedömarna värderar användningen *"av svåra ord utan att förklara dem"* i samtliga fall (n=96) som helt meningslös. När det gäller item nr 14 finns det, med hänsyn till gränsvärdet $s < .50$, inte heller någon spridning som är av betydelse. Bedömarna tycks således överensstämma väl i att det är väsentligt att fpp har faktakunskaper.

Sammanfattningsvis kan sägas, att överensstämmelsegraden i bedömarnas värdering med hänsyn till de enskilda items inom ego-ego-relationen ej är helt tillfredsställande, eftersom ca 30 % visar sig vara oreliabla.

19.2 Ego-elev-relation

Tabell 44. Ego-elev-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a_1)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
23	Förklaringar och beskrivningar	.55	.32	.72
..				
25	Stöd till eleverna	.48	.23	.68
26	Icke verbalt kontakttagande (nicka)	.57	.52	.82
27	Icke verbalt kontakttagande (peka)	.69	.51	.81
28	Användning av bekräftelser som mhm	.00		
29	Tilltal utan ögonkontakt	.36	.08	.59
30	Tal till hela klassen	.00		
31	Avbrytning av elevtal	.63	.42	.77
32	Förmåga till distanshållande	.29	.01	.53
..				
34	Favorisering av enskilda elever	.42	.15	.63
35	Att få eleverna till att arbeta	.39	.27	.70
..				
37	Uppmärksamhet riktad mot passiva elever	.49	.24	.68
38	Kontakt mellan lk och elev	.65	.45	.79
39	Rörigt i klassen	.83	.76	.90
40	Elevernas samtalsdisciplin	.54	.30	.72
41	Elevernas koncentration	.56	.33	.73
42	Självständigt arbete (elever)	.48	.23	.68
43	Elevernas inferensförmåga	.35	.07	.58
44	Skattning av elevernas förkunskaper	.55	.33	.72
45	Frågeteknik: ifyllnadsfrågor	.24	.00	.49
46	Frågeteknik: ovidkommande frågor	.34	.06	.57
47	Frågeteknik: opreciserade frågor	.23	.00	.49
48	Frågeteknik: svåra frågor	.17	.00	.43
49	Elevernas ovidkommande sysselsättning	.76	.60	.86

Kritiskt värde ($n = 90$, $\alpha = .01$): $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 44 har bedömarnas perception (a_1) för 2 items resulterat i nollkorrelationer. Därutöver syns punktskattningen för 3 items ha resulterat i otillfredsställande korrelationer. De items som visar en nedre intervallgräns, som ej uppfyller kriteriet ($r_n \geq .10$), är nr 29, 32 och 34. Således redovisas 8 items för närmare granskning i tabell 45.

Tabell 45. Ego-elev-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, perception (a_1)

Item nr	Innebörd	ml	m	s	±
28	Användning av bekräf-	1	4.48	.92	–
	telser som mhm	2	4.51	.90	
29	Tilltal utan ögonkontakt	1	6.20	.59	–
		2	6.11	.71	
30	Tal till hela klassen	1	5.47	.48	+
		2	5.54	.48	
32	Förmåga till distans-	1	6.62	.57	–
	hållande	2	6.47	.76	
34	Favorisering av en-	1	6.78	.44	+
	skilda elever	2	6.81	.37	
45	Frågeteknik: ifyllnads-	1	6.57	.73	–
	frågor	2	6.63	.68	
47	Frågeteknik: oprecise-	1	6.52	.70	–
	rade frågor	2	6.73	.57	
48	Frågeteknik: svåra	1	6.16	.76	–
	frågor	2	6.38	.56	

ml: mikrolektion s: standardavvikelse
m: medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 45 visar 6 av dessa items inom ego-elev-relationen s-värden som ej uppfyller kriteriet. För item nr 30 och nr 34 är däremot överensstämmelsen i perceptionen god, då sprindningen är låg. Slutligen uppfyller item nr 29 och nr 32 kriteriet för punktskattningen så att enbart 4 items kvarstår som oreliabla.

Sammanfattningsvis kan sägas att överensstämmelsen i bedömarnas perception, med hänsyn till de enskilda items inom ego-elev-relationen, bedöms som i stort sett tillfredsställande.

Tabell 46. Ego-ego-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a_2)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
23	Förklaringar och beskrivningar	.23	.00	.49
. .				
25	Stöd till eleverna	.62	.41	.77
26	Icke verbalt kontakttagande (nicka)	.48	.23	.68
27	Icke verbalt kontakttagande (peka)	.62	.41	.77
28	Användning av bekräftelser som mhm	.80	.67	.89
29	Tilltal utan ögonkontakt	.00		
30	Tal till hela klassen	.50	.25	.69
31	Avbrytning av elevtal	.65	.45	.79
32	Förmåga till distanshållande	.36	.08	.59
. .				
34	Favorisering av enskilda elever	.19	.00	.45
35	Att få eleverna till att arbeta	.00	.00	.00
. .				
37	Uppmärksamhet riktad mot passiva elever	.00	.00	.00
38	Kontakt mellan lk och elev	.39	.12	.60
39	Rörigt i klassen	.36	.08	.59
40	Elevernas samtalsdisciplin	.25	.00	.50
41	Elevernas koncentration	.39	.12	.71
42	Självständigt arbete (elever)	.24	.00	.54
43	Elevernas inferensförmåga	.04	.00	.32
44	Skattning av elevernas förkunskaper	.39	.12	.71
45	Frågeteknik: ifyllnadsfrågor	.18	.00	.45
46	Frågeteknik: ovidkommande frågor	.24	.00	.54
47	Frågeteknik: opreciserade frågor	.47	.22	.68
48	Frågeteknik: svåra frågor	.09	.00	.36
49	Elevernas ovidkommande sysselsättning	.45	.19	.65

Kritiskt värde ($n = 90, \alpha = .01$): $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 46 uppfyller 11 items ej kriteriet ($r_{21} \geq .27$) i punktskattningen. Därutöver uppfyller 2 items ej kriteriet $r_n \geq .10$, vilket innebär, att 13 av 24 items tycks vara oreliabla. I tabell 47 redovisas m- och s-värdena för dessa 13 items i syfte att möjliggöra en närmare granskning av orsaken till de låga reliabilitetsvärdena i bedömarnas värdering.

Tabell 47. Ego-elev-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a_2)

Item nr	Innebörd	ml	m	s	±
23	Förklaringar och beskrivningar	1	5.68	.44	+
		2	5.65	.42	
29	Tilltal utan ögonkontakt	1	5.70	.42	+
		2	5.74	.36	
32	Förmåga till distanshållande	1	3.70	.48	–
		2	3.63	.57	
34	Favorisering av enskilda elever	1	2.93	.23	+
		2	2.94	.26	
35	Att få eleverna till att arbeta	1	5.58	.46	+
		2	5.55	.40	
37	Uppmärksamhet riktad mot passiva elever	1	5.94	.34	+
		2	5.96	.17	
39	Rörigt i klassen	1	3.10	1.00	–
		2	3.31	1.29	
40	Elevernas samtalsdisciplin	1	4.86	.90	–
		2	5.04	.97	
42	Självständigt arbete	1	5.67	.36	+
		2	5.67	.33	
43	Elevernas inferensförmåga	1	5.41	.41	+
		2	5.49	.33	
45	Frågeteknik ifyllnadsfrågor	1	5.59	.46	+
		2	5.44	.46	
46	Frågeteknik: ovidkommande frågor	1	3.20	.59	–
		2	3.30	.54	
48	Frågeteknik: svåra frågor	1	3.27	.60	–
		2	3.03	.55	

ml: mikrolektion
m: medelvärde
s: standardavvikelse
±: reliabel/oreliabel

Som framgår ur tabell 47 visar 8 items spridningsvärden, som ej uppfyller kriteriet ($s \geq .50$), vilket innebär, att dessa betraktas som items "utan variation".

Sammanfattningsvis kan sägas, att överensstämmelsen i bedömarnas värdering med avseende på ego-elev-relationen betraktas som i stort sett tillfredsställande, eftersom enbart 3 av 27 items kvarstår som oreliabla.

19.3 Ego-IPO-relation

Tabell 48. Ego-IPO-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a_1)

Item nr	Innebörd	r_{21}	P ($r_n \leq P_I \leq r_ö$) = .99	
50	Bedömning av den egna undervisningen	.62	.41	.77
51	Graden av ITV-studions påverkan	.52	.28	.70
52	Grovplanering för lektionen	.48	.23	.68
53	Detaljplanering för lektionen	.40	.13	.62
54	Användning av hjälpmedel	.51	.26	.70
55	Användning av tavlan	.81	.68	.89
56	Taveldisposition	.29	.01	.53
57	Ämnespresentation	.51	.26	.70
58	Meddelande av fakta i undervisningen	.22	.00	.48
59	Anknytning till elevernas förkunskaper	.56	.33	.73
60	Utsvävningar vid framställning av ämnet	.11	.00	.38
61	Buller och oväsen utifrån	.15	.00	.41

Kritiskt värde ($n = 90, \alpha = .01$) : $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 48 uppfyller i samband med punktskattningen 4 items ej kriteriet. Item nr 56 (taveldisposition) baseras på $n_1 = 74$, $n_2 = 78$ (kritiskt värde $r = .30$). I tabell 49 redovisas för närmare granskning m- och s-värdena för dessa 4 items.

Tabell 49. Ego-IPO-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, perception (a_1)

Item nr	Innebörd	ml	m	s	±
56	Taveldisposition	1	4.03	.71	
		2	4.33	.87	–
58	Meddelande av fakta	1	4.50	.76	
	i undervisningen	2	4.70	.81	–
60	Utsvävningar vid framställning	1	6.45	.75	
	av ämnet	2	6.65	.52	–
61	Buller och oväsen uti-	1	6.83	.48	
	från	2	6.96	.19	+

ml: mikrolektion s: standardavvikelse
m: medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 49 har 3 av dessa items s-värden som uppfyller kriteriet, varför reliabiliteten i dessa måste betraktas som otillfredsställande. Item nr 61 har däremot s-värden som ej uppfyller kriteriet, vilket innebär, att detta betraktas som ett item ”utan variation”.

Sammanfattningsvis kan sägas, att samstämmigheten i bedömarnas perception av de enskilda items inom ego-IPO-relationen i stort sett är tillfredsställande.

Tabell 50. Ego-IPO-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a_2)

Item nr	Innebörd	r_{21}	P ($r_n \leq P_I \leq r_ö$) = .99	
50	Bedömning av den egna undervisningen	.04	.00	.32
51	Graden av ITV-studions påverkan	.46	.20	.66
52	Grovplanering för lektionen	.36	.08	.59
53	Detaljplanering för lektionen	.26	.00	.51
54	Användning av hjälpmedel	.46	.20	.66
55	Användning av tavlan	.00	.00	.00
56	Taveldisposition	.28	.00	.52
57	Ämnespresentation	.28	.00	.52
58	Meddelande av fakta i undervisningen	.30	.02	.54
59	Anknytning till elevernas förkunskaper	.41	.14	.62
60	Utsvävningar vid framställning av ämnet	.02	.00	.31
61	Buller och oväsen utifrån	.22	.00	.48

Kritiskt värde ($n = 96, \alpha = .01$) : $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 50 visar 5 items värden som ej uppfyller kriteriet i punktskattningen. Därutöver uppfyller 4 items ej kriteriet $r_n \geq .10$, vilket innebär, att sammanlagt 9 av 12 items har resulterat i till synes oreliabla värderingar. m- och s-värdena för dessa items redovisas i tabell 51.

Tabell 51. Ego-IPO-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a_2)

Item nr	Innebörd	ml	m	s	±
50	Bedömning av den egna	1	5.58	.26	+
	undervisningen	2	5.56	.31	
52	Grovplanering för lek-	1	6.12	.43	+
	tionen	2	6.12	.36	
53	Detaljplanering för	1	4.77	.29	+
	lektionen	2	4.79	.29	
55	Användning av tavlan	1	5.27	.47	+
		2	5.26	.42	
56	Taveldisposition	1	4.76	.80	–
		2	5.05	.66	
57	Ämnespresentation	1	6.24	.25	+
		2	6.25	.27	
58	Meddelande av fakta i	1	5.63	.55	–
	undervisningen	2	5.75	.46	
60	Utsvävning vid fram-	1	5.69	.43	+
	ställning av ämnet	2	5.72	.26	
61	Buller och oväsen uti-	1	4.49	.30	+
	från	2	4.52	.16	

ml: mikrolektion s: standardavvikelse
m: medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 51 är en granskning av bedömaröverensstämmelse, som enbart bygger på variationsskattningar, ej tillräcklig. Av de i tabell 51 redovisade 9 items uppfyller endast 2 items (56, 58) kravet på spridningen. Dessa items uppfyller dock kriteriet i punktskattningen.

Sammanfattningsvis kan sägas, att överensstämmelsen i bedömarnas värdering av de enskilda items inom ego-IPO-relationen är i stort sett tillfredsställande. Endast ett av de items som definierar detta variabelområde är med hänsyn till de angivna gränsvärdena oreliabelt.

19.4 Elev-ego-relation

Tabell 52. Elev-ego-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a_1)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
62	Negativa elevkommen-tarer	.00		
63	Elevanmärkningar på lk uppträdande	.00		
64	Elevkommentarer till lk klädsel	.00		
65	Åtlydande av lk anvis-ningar	.59	.36	.75
66	Härmningar från elever	.14	.00	.41
67	Grimaser utförda av elever	.00		
68	Motsägelser från elever	.64	.44	.78
69	Elever ställer frågor rörande ämnet	.46	.20	.66
70	Elever ställer frågor utanför ämnet	.32	.04	.55
71	Elever ger oväntade svar	.27	.00	.52

Kritiskt värde ($n = 90$, $\alpha = .01$): $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 52 uppfyller enbart 5 av 10 items kriteriet i punktskattningen. Därutöver uppfyller item nr 71 ej kriteriet i intervallskattningen. För dessa sex till synes oreliabla items redovisas m- och s-värdena för närmare granskning i tabell 53.

Tabell 53. Elev-ego-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, perception (a_1)

Item nr	Innebörd	ml	m	s	±
62	Negativa elevkommen-	1	6.99	.05	+
	tarer	2	6.98	.08	
63	Elevanmärkningar på	1	6.97	.15	+
	lk uppträdande	2	6.95	.19	
64	Elevkommentarer till	1	6.94	.37	+
	lk klädsel	2	7.00	.00	
66	Härmningar från elever	1	6.98	.11	+
		2	6.95	.24	
67	Grimaser utförda av	1	6.98	.11	+
	elever	2	6.96	.16	
70	Elever ställer frågor	1	6.77	.40	–
	utanför ämnet	2	6.69	.51	
71	Elever ger oväntade	1	5.83	.45	–
	svar	2	5.67	.59	

ml: mikrolektion s: standardavvikelse
m: medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 53 är det enbart 2 items (70, 71) som ur reliabilitetssynpunkt är tvivelaktiga fastän de uppfyller kriteriet i punktskattningen. Alla övriga visar s-värden, som ej uppfyller kriteriet, varför dessa betraktas som items "utan variation". För item nr 62 och 63 är överensstämmelsen nära nog perfekt, medan item nr 64 för ml_2 visar perfekt överensstämmelse.

Sammanfattningsvis kan sägas att överensstämmelsen i bedömarnas perception av elev-ego relationen för alla 10 items kan betraktas som tillfredsställande.

Tabell 54. Elev-ego-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a_2)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
62	Negativa elevkommen-tarer	.54	.30	.72
63	Elevanmärkningar på lk uppträdande	.62	.41	.77
64	Elevkommentarer till lk klädsel	.01	.00	.29
65	Åtlydande av lk anvis-ningar	.20	.00	.46
66	Härmningar från elever	.07	.00	.35
67	Grimaser utförda av elever	.11	.00	.40
68	Motsägelser från elever	.26	.00	.51
69	Elever ställer frågor rörande ämnet	.30	.02	.60
70	Elever ställer frågor utanför ämnet	.24	.00	.49
71	Elever ger oväntade svar	.11	.00	.38

Kritiskt värde ($n = 90$, $\alpha = .01$) : $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 54 uppfyller enbart 2 items (62, 63) både kriteriet i punktskattningen och kriteriet $r_n \geq .10$. De övriga 8 items tycks ha medfört en otillfredsställande överensstämmelse i bedömarnas värdering. För närmare granskning redovisas m- och s-värdena för dessa 8 items i tabell 55.

Tabell 55. Elev-ego-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a2)

Item nr	Innebörd	ml	m	s	±
64	Elevkommentarer till lk klädsel	1	5.50	.31	+
		2	5.48	.32	
65	Åtlydande av lk anvisningar	1	3.02	.45	+
		2	2.89	.44	
66	Härmningar från elever	1	4.03	.26	+
		2	4.02	.23	
67	Grimaser (utförda av elever)	1	3.98	.19	+
		2	3.98	.23	
68	Motsägelser från elever	1	5.14	.58	–
		2	4.97	.72	
69	Elever ställer frågor rörande ämnet	1	5.21	.42	+
		2	5.00	.45	
70	Elever ställer frågor utanför ämnet	1	5.49	.34	+
		2	5.49	.30	
71	Elever ger oväntade svar	1	2.53	.56	–
		2	2.70	.78	

ml: mikrolektion
m: medelvärde
s: standardavvikelse
±: reliabel/oreliabel

Som framgår ur tabell 55 visar 6 items s-värden, som ej uppfyller kriteriet $s \geq .50$. Enbart 2 items (68, 71) visar en otillfredsställande överensstämmelse i bedömarnas värdering.

Sammanfattningsvis kan sägas, att bedömarnas överensstämmelse i värderingen av de enskilda beteendeaspekterna inom elev-ego relationen är i stort sett tillfredsställande.

19.5 Elev-elev-relation

Tabell 56. Elev-elev-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a1)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
72	Eleverna talar i munnen på varandra	.85	.74	.91
73	Samtalar med varandra utanför ämnet	.60	.38	.76
74	Leker med varandra	.78	.64	.87
75	Diskuterar ämnet	.74	.66	.85

Kritiskt värde ($n = 90$, $\alpha = .01$) : $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 56 visar de enskilda items ett mycket gott samband i bedömarnas perception. Någon ytterligare granskning behövs ej, eftersom både kriteriet för punktskattningen och kriteriet för intraklasskorrelationens nedre gräns är uppfyllt.

Tabell 57. Elev-elev-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a2)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
72	Eleverna talar i munnen på varandra	.35	.07	.58
73	Samtalar med varandra utanför ämnet	.28	.00	.52
74	Leker med varandra	.34	.06	.57
75	Diskuterar ämnet	.00		

Kritiskt värde ($n = 90$, $\alpha = .01$) : $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 57 uppfyller item nr 75 ej kriteriet för punktskattningen, medan de övriga 3 items ej uppfyller kriteriet $r_n \geq .10$. För närmare granskning av orsaken till detta resultatutfall redovisas m- och s-värdena för de 4 items i tabell 58.

Tabell 58. Elev-elev-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a_2)

Item nr	Innebörd	ml	m	s	±
72	Eleverna talar i munnen på varandra	1	4.82	.87	–
		2	4.61	1.17	
73	Samtalar med varandra utanför ämnet	1	3.44	.63	–
		2	3.70	.82	
74	Leker med varandra	1	4.76	.79	–
		2	4.57	1.00	
75	Diskuterar ämnet	1	4.48	.47	–
		2	4.31	.57	

ml: mikrolektion
m: medelvärde
s: standardavvikelse
±: reliabel/oreliabel

Som framgår ur tabell 58 uppfyller samtliga s-värden kriteriet, dvs alla items visar variation.

Bedöms sambandet enbart med hänsyn till kriteriet i punktskattningen kan sägas, att överensstämmelsen i bedömarnas värdering inom elev-elev-relationen är i huvudsak tillfredsställande. Visserligen är dessa korrelationer, med undantag av item nr 75, enligt punktskattningen signifikanta, men intervallskattningen visar, att kriteriet för r_n ej är uppfyllt, dvs man skulle kunna få mycket låga korrelationer.

Sammanfattningsvis kan sägas att överensstämmelsen i bedömarnas värdering betraktas som i stort sett tillfredsställande.

19.6 Elev-IPO-relation

Tabell 59. Elev-IPO-relation: Intraklasskorrelationer, punkt- och intervallskattningar, perception (a_1)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
76	Elevernas intresse	.56	.33	.73
77	Framställning av ämnet	.57	.34	.74
78	Elevernas reaktion inför ämnet	.57	.34	.74
79	ITV-studions påverkan på eleverna	.34	.06	.57

Kritiskt värde ($n = 90$, $\alpha = .01$): $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 59 uppfyller samtliga items kriteriet i punktskattningen. Med hänsyn till kriteriet för r_n bör enbart item nr 79 granskas närmare. Detta item har följande m- och s-värden: ml_1: m 4.88, s .83 och ml_2: m 5.28, s .74.

Spridningen i item nr 79 uppfyller kriteriet.

Sammanfattningsvis kan sägas, att inget av dessa 4 items i den aktuella analysen behöver betraktas som oreliabelt, även om item nr 79 ej är helt tillfredsställande.

Tabell 60. Elev-IPO-relation: Intraklasskorrelationer, punkt- och intervallskattningar, värdering (a_2)

Item nr	Innebörd	r_{21}	$P(r_n \leq P_I \leq r_ö) = .99$	
76	Elevernas intresse	.48	.23	.68
77	Framställning av ämnet	.50	.25	.69
78	Elevernas reaktion inför ämnet	.09	.00	.36
79	ITV-studions påverkan på eleverna	.27	.00	.52

Kritiskt värde ($n = 90$, $\alpha = .01$): $r = .267$, Fisher, 1946, s 209

Som framgår ur tabell 60 uppfyller ej item nr 78 kriteriet för punktskattningen. Med hänsyn till kriteriet för intervallskattningens nedre gräns är därutöver item nr 79 tvivelaktigt ur reliabilitetssynpunkt. I tabell 61 redovisas m- och s-värdena för dessa 2 items (78, 79).

Tabell 61. Elev-IPO-relation: Items som ej uppfyller kriteriet $r_{21} \geq .27$ eller $r_n \geq .10$, värdering (a_2)

Item nr	Innebörd	ml	m	s	±
78	Elevernas reaktion inför	1	5.42	.65	
	ämnet	2	5.51	.61	–
79	ITV-studions påverkan	1	3.16	.81	
	på eleverna	2	2.93	.79	–

ml: mikrolektion s: standardavvikelse
m: medelvärde ±: reliabel/oreliabel

Som framgår ur tabell 61 beror det låga sambandet ej på att spridning saknas. Item nr 78 bör betraktas som oreliabelt, medan item nr 79 tycks vara tvivelaktigt ur reliabilitetssynpunkt, fastän detta item uppfyller kriteriet i punktskattningen.

Sammanfattningsvis kan sägas, att överensstämmelsen i bedömarnas värdering med hänsyn till de enskilda items inom elev-IPO-relationen i huvudsak kan betraktas som tillfredsställande.

19.7 Sammanfattning

Reliabiliteten i bedömarnas skattningar granskades ur flera olika aspekter. Efter beräkningen av intraklasskorrelationen (r_{21}) med hjälp av en variansanalytisk design utfördes signifikansprövningar. Men även intervallskattningar har redovisats och diskuterats, för att bestämma i vad mån korrelationerna påvisbart skiljer sig från noll, respektive för att bestämma de gränser, inom vilka man kan förvänta sig att korrelationerna kommer att ligga. En granskning av orsaken till bristande reliabilitet hos items medförde, att avgörandet avseende ett visst items reliabilitet ej enbart kunde baseras på dessa mått. Med utgångspunkt från kriteriet $s < .50$ för en godtagbar spridning (bedömningen är subjektiv), granskades de items som ej uppfyllde dels kriteriet $r_{21} \geq .27$, dels kriteriet $r_n \geq .10$. Denna granskning visade, att det finns ett antal items "utan variation". I sådana fall blir ett reliabilitetsmått, baserat på varianser, en indikator, som ej går att använda på överensstämmelsen i bedömarnas skattningar. Att det ej finns någon variation (kriteriet $s < .50$) tolkas i denna analys som att överensstämmelsen i bedömarnas skattningar är nära nog perfekt. Men detta betyder samtidigt, att bedömningar av items "utan variation" har utförts på ett klichéartat sätt, dvs lika för alla eller nästan alla fpp.

Med utgångspunkt från de enskilda items bedömdes reliabiliten inom respektive variabelområde som tillfredsställande när *alla* items uppfyllde antingen kriteriet $r_{21} \geq .27$ eller $s < .50$. I de fall, där ett eller flera items ej uppfyllde dessa kriterier, bedömdes reliabiliteten som i stort sett tillfredsställande. Resultatet av reliabilitetsgranskningen sammanfattas i tabell 62.

Tabell 62. Sammanfattning av antalet reliabla/oreliabla items för variabelområdena 1–6

Variabel-område		antal totalt	perception (a_1) +	perception (a_1) –	värdering (a_2) +	värdering (a_2) –
1	Ego-ego	20	17	3	14	6
2	Ego-elev	24	20	4	21	3
3	Ego-IPO	12	9	3	10	2
4	Elev-ego	10	10	0	8	2
5	Elev-elev	4	4	0	3	1
6	Elev-IPO	4	4	0	3	1
	S:a	74	64	10	59	15

+: reliabla items
–: oreliabla items

Av de 74 items som ingått i analysen är det för perception (a_1) ca 13 % och för värdering (a_2) ca 20 %, som måste betraktas som oreliabla. Med hänsyn till detta resultat har bedömarnas observationsdata för den fortsatta resultatanalysen sammanslagits.

Mot bakgrund av resultatet som framgår ur tabell 62 kan skattnings- och värderingsformulär F III betraktas som i stort sett tillfredställande ur reliabilitetssynpunkt.

20. Reliabiliteten i summavariablerna

På samma sätt som i resultatanalys 1 (del 2) baseras nivåanalyserna, utförda på bedömarnas medelbedömning, ej på varje enskilt item, utan på variabelområdena. Trots den ingående reliabilitetsdiskussionen i kapitel 19 ovan kan det ha ett visst intresse att försöka skatta reliabiliteten i respektive summavariabel. Ett vanligt tillvägagångssätt är, att man skattar sambandet mellan de enskilda items, som utgör ett test eller deltest, med hjälp av Spearman—Browns "prophecy"-formel (Rozeboom, 1966, s 409). För att kunna tillämpa denna formel, måste man dock kunna förutsätta *lika* mellan-item-korrelation, en förutsättning som i praktiken aldrig kan uppfyllas när $n > 2$, i de flesta fall inte ens approximativt. Att välja en sådan korrelation framför en annan för skattningen av summavariabelns reliabilitet (r_s) skulle medföra dels att större delen av den tillgängliga homogeniteten ignoreras, dels att ett arbiträrt val införs, ett element som ej avsågs ingå i formeln. Denna svårighet skulle ha kunnat undvikas med hjälp av en genomsnittlig mellan-item-korrelation för r_{ii}. Av praktiska skäl är det av mindre intresse hur man under användning av Spearman—Browns formel skattar testets reliabilitet, eftersom Cronbachs "coefficient alpha" (α_C) är ett bättre tillvägagångssätt (se Rozeboom, 1966, ss 408—415).

α_C är en skattning, som i första hand är baserad på en variabelgrupp, men den kan också beräknas på olika summavariabler. Används den på summavariabler, blir koefficientens korrelationsvärde beroende av hur man analyserar delar av ett test (t ex delskalor). α_C för hela testet behöver alltså ej nödvändigtvis vara identiskt med α_C för deltesten.

De problem, som uppstått vid bestämningen av på vilket sätt man skulle kunna använda homogeniteten i ett test(testets interna konsistens) som reliabilitetsskattning, ledde till α_C. Definieras homogeniteten för ett antal items som kvot av den genomsnittliga mellan-item-kovariansen och testets genomsnittliga varians så är α_C lika med Spearman—Browns formel, dock med undantaget att r_s och r_{ii} har blivit ersatta med α_i och homogeniteten (h_i) (se Rozeboom, 1966, 412).

Innehållet i h_i som en mätning av mellan-item-ekvivalens är, det bör poängteras, en approximering av den genomsnittliga mellan-item-korrelationen och en exakt mätning *endast om varje enskilt item* har samma varians. α_{Ci} för en

summavariabel är alltså en generalisering av Spearman—Browns formel (förutsätter identiska item-korrelationer och interkorrelationer) och kan i motsats till denna beräknas för vilken arbiträr analys av test eller deltest som helst.

Som redan har diskuterats i del 2 kan man betrakta testets interna konsistens (homogenitet), såsom den skattas genom α_C, som en nedre gräns för testets reliabilitet. Detta gäller dock enbart så länge mätfelen mellan de enskilda items är okorrelerade. Rozeboom (1966, s 415) skriver: "Distressingly, however, this is a most unrealistic assumption." Dessutom bör man ej glömma, att α_C är en funktion av homogenitet *och* antal items i ett visst test. Nedan följer en redovisning av α_C för bedömarnas observationsdata. I tabellerna 63 och 64 redovisas dessa koefficienter för perception (a_1) och värdering (a_2). Inom parentes anges lkk α_C för att möjliggöra en jämförelse mellan båda datauppsättningarna. Dessutom anges homogeniteten (r_s), beräknad med Spearman—Browns formel (med alla förbehåll som nämndes ovan). För redovisning av r_{ii} har tabellernas högra kolumn använts.

Tabell 63. Reliabilitet för summavariabel, alfa-koefficient (α_C), homogenitet (r_s) och medelreliabilitet, baserad på $\bar{r}_{21}$, perception (a_1)

Variabel-område		α_C ml$_1$ k = 2	n = 96	ml$_2$ k = 2	n = 96	r_s	$\bar{r}_{21}$
1	Ego-ego	.73	(.72)	.73	(.79)	.94	.44
2	Ego-elev	.66	(.70)	.57	(.57)	.95	.44
3	Ego-IPO	.32	(.33)	.39	(.35)	.90	.43
4	Elev-ego	.20	(.57)	.52	(.64)	.76	.24
5	Elev-elev	.00	(.00)	.08	(.31)	.91	.74
6	Elev-IPO	.83	(.64)	.79	(.53)	.82	.54

k = 2 : bedömare, n = 96 : lärarkandidater, ml : mikrolektion

Tabell 64. Reliabilitet för summavariabel, alfa-koefficient (α_C), homogenitet (r_s) och medelreliabilitet, baserad på $\bar{r}_{21}$, värdering (a_2)

Variabel-område		α_C ml$_1$ k = 2	n = 96	ml$_2$ k = 2	n = 96	r_s	$\bar{r}_{21}$
1	Ego-ego	.05	(.00)	.19	(.12)	.92	.37
2	Ego-elev	.24	(.16)	.03	(.12)	.92	.33
3	Ego-IPO	.19	(.37)	.04	(.43)	.80	.26
4	Elev-ego	.31	(.74)	.19	(.80)	.77	.25
5	Elev-elev	.00	(.00)	.11	(.00)	.56	.24
6	Elev-IPO	.21	(.31)	.04	(.24)	.67	.34

k = 2 : bedömare, n = 96 : lärarkandidater, ml : mikrolektion

En jämförelse mellan tabellerna 63 och 64 implicerar, att α_C som mått på reliabiliteten varit mera adekvat i samband med skattningen av reliabiliteten i bedömarnas perception (a_1) än vid skattningen av reliabiliteten i bedömarnas värdering (a_2).

När det gäller perceptionen visar α_C inom ego-dimensionen ungefär samma värden för både bedömare och lkk. För variabelområdena 1 och 2 indikerar α_C ett tillfredsställande reliabilitetsvärde, ett resultat som sammanfaller med bedömningen av reliabiliteten utförd på itemnivå. Ett mycket högt värde registreras för variabelområde 6. Lkk skiljer sig dock här mera markant från bedömarna.

Av särskilt intresse är frågan varför α_C för variabelområdena 3, 4 och 5 har resulterat i låga värden. Det låga värdet för variabelområde 3 kan förklaras med att de 12 items som operationellt beskriver denna summavariabel har en mycket låg summavarians. Det finns tydligen mycket liten mellan-item varians.

För variabelområde 4 måste man söka orsaken till den obetydliga korrelationen i det faktum, att 6 av de 10 items ej har någon varians alls, dvs perceptionen är stereotyp. I detta fall kan α_C ej fungera, då den förutsätter varians. En granskning av variansen för de enskilda items visar, att bedömarna ändrade sin inställning vid undervisningstillfälle 2, eftersom det plötsligt finns varians för flera items. Resultatet reflekteras i en α_C, som återspeglar samma bedömning hos experterna som hos lkk.

För variabelområde 5 syns α_C återspegla ett realistiskt förhållande. Det finns i detta fall en stor varians för varje enskilt item, och itemvariansen överstiger summavariansen (α_C blir negativ). Detta resultat implicerar att summavariabeln är oreliabel. Som framgår ur tabell 63 är dock homogeniteten hög. Det kan vara till hjälp att här upprepa, att α_{Ci} går mot 1 när n (items) stiger. Hastigheten med vilken denna punkt nås avtar dock med avtagande homogenitet. Denna asymptotiska funktion av α_C i förhållande till 1 bör kunna klargöra, att α_C ej är en mätning av itemhomogenitet. Homogenitet inom ett antal items bör skattas på annat sätt (se t ex Rozeboom, 1966, s 321). I bedömarnas värdering (a_2) finns det till synes ej någon reliabilitet alls. Som framgått av reliabilitetsdiskussionen (kap 19) finns det emellertid för denna aspekt ett relativt stort antal items där bedömarnas skattning är klichéartad. α_C kan alltså, på grund av att det finns ett relativt stort antal items, där sambandet visserligen är nära nog perfekt men utan variation, ej vara någon lämplig skattning av reliabiliteten i summavariablerna. En något större variabilitet finns i lkk värdering, där α_C mera korrekt återspeglar en låg reliabilitet för summavariablerna. Däremot finns inom variabelområde 4 en markant skillnad mellan α_C-värdena för lkk och bedömarnas värdering. En av orsakerna till detta resultat skulle kunna vara att lkk värdering visar en jämförelsevis mycket större variabilitet än bedömarnas värdering. För lkk, i motsats till bedömarna, gäller såväl inom-item, som mellan-item-varians.

Sammanfattningsvis kan sägas, att instrumentets reliabilitet bedöms som i stort sett tillfredsställande. Den utförliga reliabilitetsdiskussionen har förts för att belysa precisionen i skattningarna och för att undvika en förhastad bedömning av tillförlitligheten i bedömarnas perception och värdering. Sambandet i dessa observationsdata betraktas som gott (i ett antal fall är det "perfekt"). Som tidigare har anförts, är syftet med expertbedömningarna att erhålla ett objektivitetsmått. För att kunna granska i vilken utsträckning undervisningsprocessen återspeglas på ett objektivt eller realistiskt sätt i lkk informationsselektion, genererades en "medelbedömning" ur bedömarnas observationsdata.

Syftet med den nedan följande nivåanalysen, baserad på bedömarnas medelbedömning, är att granska, om det i bedömarnas perception och värdering finns skillnader, som hänför sig till de olika experimentella påverkningarna.

21. ANOVA av bedömarnas medelbedömning

Vid granskning av eventuella nivåskillnader i de pedagogiska experternas perception och värdering av lkk undervisningsbeteende har ANOVA-tabeller som innehöll exakt samma variationsorsaker som ANOVA-tabellerna vid analysen av lkk observationsdata kommit till användning.

Genom (1) ett yttre kriterium, definierat som bedömarnas medelbedömning och (2) ett konstanthållande av observationsobjektet, dvs experimentets videobandade mikro-lektioner, blir det möjligt att studera, om det finns observerbara skillnader i beteende mellan olika grupper av lkk. Finns det skillnader i testvärdena, måttet på den återkopplade informationen, som är signifikanta på den valda nivån, innebär detta att bedömarna har selegerat skilda undervisningsbeteenden, som hänför sig till experimentets olika påverkningsgrupper. Analys- och redovisningstekniken är densamma som har tillämpats i del 2. Även hypotesprövningen avser samma hypoteser, nämligen:

H_0: traditionell handledning ger ingen effekt
H_0: självkonfrontation via ITV/VR ger ingen effekt
H_0: samspel mellan traditionell handledning och självkonfrontation via ITV/VR föreligger ej.

21.1 Mönstret i F-testen

I detta avsnitt studeras närmare i vilken utsträckning det finns ett mönster i F-testen. En sammanfattning av de signifikanta effekterna i de enskilda ANOVA-tabellerna (se Bierschenk, 1972 e, grön bilaga 2:1—2:6) redovisas i tabell 65. (Om ej något annat anges, refereras i del 3 till dessa bilagor.) I diskussionen av de enskilda variationsorsakerna utelämnas huvudeffekterna i faktorerna A och V samt alla samspelseffekter med faktor V som variationsorsak (för diskussion hänvisas till del 1). I samband med diskussionen av de nedan följande resultaten relateras utfallet till del 2, men en signifikansprövning av skillnaderna och en närmare granskning sker först i resultatanalys 3 (se del 4).

Faktor T har ej lett till signifikanta huvudeffekter, varför nollhypotesen accepteras. Med bedömarnas medelbedömning som jämförelsekriterium innebär detta resultat, att den externt förmedlade självkonfrontationen via ITV/VR-tekniken ej har medfört någon påvisbar skillnad i lkk beteenden.

Tabell 65. Bedömarnas medelbedömning: Sammanfattning av ANOVA-tabellerna

Variations-orsak	Variabelområde 1	2	3	4	5	6
T						
H				**	**	
TH	**	**		*		**
U		*	**	**	**	
UT				*		
UH				**	*	
UTH						
A	**	**	**	**	**	**
AT						
AH				**	**	
ATH	**	**	**	**		**
AU						
AUT						
AUH			*			
AUTH						
V	**	**	**	**	**	**
TV						
HV				*	**	
THV	**	**	**	**		
UV			**		**	
UTV					*	
UHV					*	
UTHV						
AV	**	**	**	**	**	**
ATV						
AHV		**		*	**	*
ATHV	**	**	**	**		**
AUV		*	**		**	**
AUTV			*			
AUHV				*	*	
AUTHV			**			

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
V: Skattnings- och värderingsformulär F III
** $F_{.99}(1, 92) = 7.08$
* $F_{.95}(1, 92) = 4.00$

Faktor H har lett till två signifikanta huvudeffekter (4, 5). När det gäller dessa variabelområden förkastas nollhypotesen. Testat mot medelbedömningen som kriterium finns det påvisbara skillnader i protokollmaterialet som en följd av handledarens kommentarer, dels avseende elevernas beteenden riktade mot lk (4), dels avseende elevernas beteenden sinsemellan (5). I kap 5.1.2 har handledningen under experimentet beskrivits. I detta sammanhang anfördes bl a, att handledaren koncentrerade sin handledning på problem som t ex elevernas uppmärksamhet, kunskap eller aktivitet. Medelbedömningen indikerar, att lkk undervisningsbeteenden till följd av handledarens kommentarer påvisbart skiljer sig från de övriga lkk beteenden.

Som resultatanalysen i del 2 visar, kan någon huvudeffekt i faktor H ej påvisas i fråga om lkk självbedömning. Men resultatet i föreliggande analys implicerar enligt bedömarna, att handledaren framgångsrikt har förmedlat dels sin undervisningsstrategi, dels konkreta handlingsförslag och att lkk tycks ha lyckats modifiera sitt eget beteende på basis av handledningsdiskussionen. Denna effekt har dock ej kunnat observeras av lkk själva. Lkk perceptions- och värderingsstruktur synes ej ha blivit påverkad av handledaren. Enligt den ovan förda diskussionen synes lkk perception och värdering ej återspegla den realistiska undervisningssituationen, dvs lkk diskrimineringsfärdighet tycks ännu inte tillåta, att de identifierar och klassificerar beteenden enligt pedagogisk-psykologiska normer (jfr del 1).

Faktor U har lett till påvisbara effekter inom variabelområdena 2, 3, 4 och 5. Oavsett påverkan finns det förändringar i lkk undervisningsbeteenden från tillfälle 1 till 2. Nollhypotesen förkastas för denna faktor avseende dessa variabelområden. I en jämförelse mellan detta resultatutfall och lkk självbedömning överensstämmer observationen av förändringar i två fall (2, 3). Förändringar enligt bedömarnas medelbedömning finns dessutom för variabelområdena 4 och 5, medan lkk perception och värdering indikerar förändringar i variabelområde 1 (ego-ego-relationen). Både experterna och lkk tycks överensstämma i bedömningen att det ej förekommit någon förändring inom variabelområde 6.

Påvisbara effekter i faktorkombinationen TH finns i variabelområdena 1, 2, 4 och 6. För dessa variabelområden förkastas nollhypotesen. Någon sådan effekt finns däremot ej för variabelområdena 3 och 5. Detta resultat implicerar, att det finns påvisbara skillnader, oavsett om det gäller perception eller värdering, i lkk undervisningsbeteenden som följd av en kombination av traditionell handledning *och* externt förmedlad självkonfrontation via ITV/VR. Någon liknande effekt kunde i samband med lkk observationer enbart påvisas för variabelområdena 2 och 4. En uppföljning av faktorkombinationen TH med hänsyn till precisionsfaktor A visar, att det (med medelbedömningen som testkriterium) finns påvisbara effekter i alla variabelområdena, utom 5. Alla effekter är signifikanta med $\alpha = .01$. En jämförelse mellan detta resultat och resultatanalys 1 visar, att även lkk självbedömningar har resulterat i påvisbara

ATH-effekter, men med undantag av variabelområde 6 är dessa signifikanta med $\alpha = .05$. En kolumnvis genomförd granskning av bedömarnas medelbedömning visar, förutom de redan diskuterade effekterna, för variabelområde 3 en effekt i AUH-kombinationen. Lkk bedömningar har i stället resulterat i en AUT-effekt. AUH-samspelet kan i samband med de övriga effekterna inom variabelområde 3 ses som ett tolkbart resultat, men jämfört med de övriga variabelområdena är detta resultat isolerat och bör ej tillmätas någon större vikt. Skillnaden mellan bedömarnas och lkk observationer avseende denna grupp av variationsorsaker ligger dessutom i att det i samband med medelbedömningen förekommer enbart en påvisbar effekt ($\alpha = .05$), medan det i resultatanalys 1 finns sex påvisbara effekter.

En granskning av faktorerna T och H kolumnvis (4 och 5) visar, att faktor T endast har resulterat i en påvisbar samspelseffekt (UT (4)). Faktor H har gett ett kraftigare utslag för båda variabelområdena. Det finns påvisbara effekter i samspelen UH och AH. Detta resultat överensstämmer ej helt och hållet med de effekter som redovisades för lkk observationer. AH-samspelet visar i samband med lkk observationer påvisbara effekter för variabelområdena 5 och 6, medan UH-samspel över huvud taget ej förekommer. Slutligen bör nämnas, att det i samband med bedömarnas medelbedömning förekommer avsevärt flera signifikanser i de faktorkombinationer som involverar faktor V än i resultatanalysen av lkk observationer, vilket tyder på att bedömarnas perception är mera differentierad än lkk.

Granskningen av mönstren i F-testen visar, att de finns två påvisbara huvudeffekter i faktor H, medan faktor T ej har givit något utslag. TH-samspelen har i 4 av de 6 variabelområdena resulterat i påvisbara effekter, och en uppföljning av TH-kombinationen med hänsyn till precisionsfaktorn A visar i alla variabelområden (med undantag för område 5) påvisbara effekter.

Granskas faktorerna T och H kolumnvis, framträder ett regelbundet mönster för variabelområdena 4 och 5. I båda fallen är enbart faktor H involverad i faktorkombinationen. Mot bakgrund av det diskuterade mönstret finns det dessutom två relativt isolerade signifikanser (UT (4), AUH (3)), som i relation till respektive kolumn kan vara av intresse men som, sedda över hela tablån, synes vara av mindre betydelse. I diskussionen påvisades på vilka punkter lkk självbedömning (resultatanalys 1) och bedömarnas medelbedömning (resultatanalys 2) överensstämmer eller ej. Överlag är mönstret i F-testen, som är baserade på medelbedömningen, mera enhetligt än mönstret i de F-test som har redovisats i resultatanalys 1.

21.2 Precisionen och styrkan i F-testen

Innan någon mera detaljerad analys av de signifikanta F-testen redovisas, granskas nedan, mot bakgrund av den argumentation som fördes i del 2, pre-

cisionen och styrkan i F-kvoterna. Proportionen av varians (PV) som förklaras av de enskilda faktorkombinationerna redovisas för de signifikanta F-testen i tabell 66.

Tabell 66. $\hat{\omega}^2$-värden för signifikanta F-kvoter: summavariabel

Variations-orsak	Variabelområde 1	2	3	4	5	6
T						
H				.00	.03	
TH	.02	.01		(.00)		.02
U		(.00)	.01	.00	.01	
UT				(.00)		
UH				.00	(.01)	
UTH						
A	.33	.53	.63	.96	.36	.02
AT						
AH				.00	.01	
ATH	.01	.02	.02	.00		.04
AU						
AUT						
AUH			(.00)			
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
(): Anger $F_{.95}$ (1, 92) = 4.00

Som framgår ur tabell 66 ligger proportionen systematisk varians, som förklaras av de enskilda faktorerna och faktorkombinationerna (med undantag av faktor A) mellan .00 och .04. ω^2 har tolkats som proportion-varians (PV) som hänför sig till alternativa effekter. I en faktoriell design är ω^2 en funktion av antalet effekter som prövas, dvs ω^2 ger uttryck för PV som ett resultat av m_1 och ej av m. Ett mått på storleken av en enskild effekt, när alla övriga hålles konstanta, är Cohens f. Både skattningen av f och sannolikheten (g), dvs att en enskild effekt är av denna storleksordning, anges i tabell 67.

Styrkan i precisionsvärdena är i ett flertal fall mycket god. Sannolikheten för att man skall kunna upptäcka sådana skillnader är mycket högre än i resultatanalys 1. Tolkningen av dessa resultat torde därför vara mera övertygande, dvs resultaten återspeglar i hög grad ett realistiskt förhållande. Styrkeskattningen implicerar för denna analys, att PV i ett flertal fall verkligen är av den storleksordning som har redovisats i tabell 66 ovan. Den systematiska variansen mellan faktor A (aspekter) och den beroende variabeln visar, utom för elev-

Tabell 67. Effektstorlek och dennas sannolikhet för bedömarnas medelbedömning

Variations-orsak	Variabelområde 1	2	3	4	5	6
1. Effektstorlek (f)						
T						
H				.17	.18	
TH	.21	.21		(.14)		.19
U		(.12)	.16	.19	.14	
UT				(.18)		
UH				.23	(.18)	
UTH						
A	.71	1.04	1.24	5.49	.99	.14
AT						
AH				.22	.28	
ATH	.31	.41	.49	.56		.33
AU						
AUT						
AUH			(.23)			
AUTH						
2. Sannolikhet (g)						
T						
H				.72	.78	
TH	.62	.62		(.33)		.52
U		(.40)	.67	.83	.46	
UT				(.46)		
UH				.77	(.46)	
UTH						
A	>.99	>.99	>.99	>.99	>.99	.54
AT						
AH				.67	.89	
ATH	.66	.91	>.99	>.99		.73
AU						
AUT						
AUH			(.55)			
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
(): Anger $F_{.95}$ (1, 92) = 4.00

IPO-relationen, genomgående mycket höga värden. Precisionen i bedömarnas användning av aspekterna synes vara större än precisionen vid tillämpningen av faktor A i samband med lkk observationer. Även styrkeskattningen visar sannolikheten > .99, dock med undantag av variabelområde 6, där sannolikheten ligger omkring .50 men troligen är mindre än så. En jämförelse mellan tabell 66 och motsvarande tabell 33 i resultatanalys 1 visar, att den systematiska variansen med hänsyn till de enskilda faktorerna och faktorkombinationerna har resulterat (med någon % skillnad) i samma värden. Som f (tab 67) indikerar har samspelseffekterna (med något undantag) resulterat i små eller mellanstora effekter.

Styrkan i samspelseffekterna skiljer sig dock avsevärt. Av 11 signifikanta samspelseffekter (med $\alpha = .01$) har endast 2 resulterat i g-värden som ligger omkring .50.

21.3 Efterprövningar

Jämfört med lkk självbedömningar har bedömarnas medelbedömning resulterat i ett mera enhetligt mönster i F-testen, i en större precision och i högre sannolikheter för de påvisade effekterna. Bedömarnas medelbedömning utgör alltså ett mera beviskraftigt tolkningsunderlag än lkk självbedömningar. Detta resultatutfall är dock ej oväntat. Mycket av variabiliteten i lkk självbedömningar måste nämligen hänföras till individskillnader, som har existerat redan före experimentets början. Eftersom analysen av lkk observationsdata bygger på $n = 96$, medan analysen av bedömarnas medelbedömning bygger på $k = 2$, blir spridningen i bedömarnas observationer mindre, vilket i sin tur har till följd att skillnaderna mellan cellmedelvärdena ej behöver vara av samma storleksordning som för lkk för att ge påvisbara effekter.

För de medelstora och stora effekterna, som uppfyller kriteriet $g > .70$ skulle i detta sammanhang en mera detaljerad analys och diskussion kunna vara befogad. Kontrastanalyserna, försedda med några kommentarer redovisas dock i enlighet med den tidigare tillämpade redovisningstekniken, i Bierschenk, 1972 e, grön bilaga 3.

22. Sammanfattning av ANOVA-resultaten

Vill man bestämma objektiviteten i lkk självbedömningar i samband med egna lektioner, behöver man dels jämföra lkk egna observationer med ett för lkk externt kriterium, dels hålla undervisningsprocessen konstant. "Objektivitet" definierades för denna undersökning som bedömaröverensstämmelse dels mellan de enskilda pedagogiska experterna, dels mellan experter och lkk.

I denna resultatanalys granskades perceptionen och värderingen av experimentets videobandade mikrolektioner utifrån expertbedömningar, som utgör det yttre kriteriet. Det torde vara allmänt accepterat, att externa observatörer kan åstadkomma större objektivitet vid bedömning av en persons beteenden än personen själv. För att man skall kunna studera lkk objektivitet, definierad genom expertbedömningar, krävs det att bedömningskategorierna för lkk och bedömarna är identiska, varför även expertbedömningar utfördes med hjälp av skattnings- och värderingsformulär F III.

I den ovan redovisade analysen granskades om och i vilken utsträckning lkk, som utsattes för olika experimentella påverkningar, skiljer sig påvisbart i de skattade undervisningsbeteendena. Som kriterium har använts bedömarnas medelbedömning. Den variansanalytiska bearbetningen av bedömarnas datauppsättning är således en spegelvänd analys i förhållande till resultatanalysen av lkk självbedömningar. ANOVA har i den aktuella analysen tillämpats på två olika sätt.

22.1 ANOVA-modellen som analysteknik i samband med reliabilitetsstudier

I kapitel 19 redovisades och diskuterades ingående intraklasskorrelationskoefficienten r_{21} (två bedömare, ett undervisningstillfälle). Precisionen i dessa skattningar granskades med hjälp av intervallskattningar. Visade ett item låg reliabilitet, granskades spridningsvärdena för respektive item närmare. I vad mån de items som har resulterat i låga reliabilitetsvärden verkligen måste betraktas som oreliabla, beror nämligen i hög grad på med vilka skattningsmetoder man försöker bestämma reliabiliteten i en mätning. En granskning av de till synes oreliabla items visade, att det ofta ej fanns någon variation (kriteriet $s < .50$). Ett reliabilitetsmått baserat på varianser blir i sådana fall en indikator som ej går att använda för att skatta överensstämmelsen i bedömarnas observationer.

Det kan sägas, att bedömarna i ett antal fall har utfört sina skattningar på ett klichéartat sätt, dvs nästan lika för alla lkk (n = 96), vilket innebär, att det finns antingen en nära nog perfekt eller en helt perfekt överensstämmelse i bedömarnas skattningar. Granskningen av reliabiliteten i de enskilda items visade ett i huvudsak tillfredsställande resultat.

I kapitel 20 granskades och diskuterades reliabiliteten i de enskilda summavariablerna (variabelområdena 1—6). För skattningen av reliabiliteten har använts Cronbachs alfa (α_C). Användbarheten av denna reliabilitetskoefficient diskuterades utförligt. α_C har dock i vissa fall varit ett olämpligt mått, eftersom vissa variabelområden resulterade dels i en mycket låg summavarians, dels i en mycket låg mellanitemvarians. Förklaringen är bl a, att det i dessa variabelområden finns ett relativt stort antal items där bedömarnas skattning varit "utan variation". α_C kan alltså, när det finns ett relativt stort antal items där sambandet visserligen är nära nog perfekt men utan variation, ej fungera som någon lämplig skattning av reliabiliteten i summavariablerna.

Som ett jämförelsekriterium redovisades i samband med α_C också ett homogenitetsmått, skattat med hjälp av Spearman—Browns formel (med alla förbehåll). Den utförliga reliabilitetsdiskussionen har förts för att belysa precisionen i skattningarna och för att undvika en förhastad bedömning av tillförlitligheten i bedömarnas perception och värdering. Överensstämmelsen i dessa observationsdata betraktas som tillfredsställande, och i ett antal fall är den perfekt.

Resultatet av reliabilitetsgranskningen ledde till konstruktionen av en medelbedömning. På basis av denna medelbedömning utfördes en ANOVA för varje enskilt variabelområde.

22.2 ANOVA-modellen som analysteknik för granskning av nivåskillnader

Sex ANOVA har utförts för att man skall kunna granska om det finns skillnader i lkk undervisningsbeteenden till följd av den experimentella påverkan. Eftersom analysen genomfördes exakt på samma sätt som vid resultatanalysen av lkk självbedömningar, gäller samma argumentation som har redovisats i del 2. De enskilda ANOVA-tabellerna har redovisats i grön bilaga 2. Granskningen utfördes såsom tidigare i tre steg.

Mönstret i F-testen visade dels flera signifikanser, dels större enhetlighet än motsvarande analys av lkk självbedömningar. Dessutom har bedömarnas medelbedömning inom elev-ego-relationen och elev-elev-relationen, avseende påverkningsformen traditionell handledning (faktor H), resulterat i huvudeffekter.

Huvudeffekterna indikerar dels när det gäller elevernas beteenden mot lk, dels elevernas beteenden sinsemellan, högre (positiva) värden för de lkk som

fick handledning. Orsakerna till detta resultat kan vara flera. En möjlig förklaring skulle kunna vara, att handledaren har påverkat lkk så, att de åstadkommit såväl ett positivare beteende hos eleverna gentemot lk som ett positivare beteende hos eleverna sinsemellan. Men effekten kan mycket väl tänkas bero på handledarens närvaro i klassen, vilken kan ha haft en dämpande effekt på elevernas beteenden. Vilken av orsakerna som är den mest troliga kan ej avgöras i detta sammanhang.

Att bedömarnas perception skiljer sig i samtliga sex variabelområden från värderingen (faktor A) framgår av mönstret för denna faktor. Av samma skäl som anfördes i den tidigare diskussionen har de effekter som hänför sig till faktorerna A och V ej granskats och tolkats. Icke signifikanta effekter i dessa faktorer hade däremot krävt en närmare granskning.

Undervisningstillfällena (faktor U) har också resulterat i påvisbara huvudeffekter. Resultatet indikerar, att det oavsett den experimentella påverkan har förekommit beteendemodifikationer hos lkk med hänsyn till ego-elev-relationen, ego-IPO-relationen, elev-ego-relationen och elev-elev-relationen. En tillräcklig styrka för en tolkning av dessa effekter kan dock enbart konstateras för elev-elev-relationen.

Signifikanta samspelseffekter i faktorkombinationerna TH och ATH kan för bedömarnas medelbedömning påvisas inom ego-ego-relationen, ego-elev-relationen, elev-ego-relationen och elev-IPO-relationen. Enligt bedömarna har alltså externt förmedlad självkonfrontation via ITV/VR i kombination med traditionell handledning medfört beteendemodifikationer oavsett om det gällt perception eller värdering. Sådana effekter påvisas i lkk självbedömning enbart inom ego-elev-relationen och elev-ego-relationen. Dessutom har medelbedömningen resulterat i medelstora och stora ATH-effekter med relativt höga styrkevärden, som kan påvisas i alla variabelområden utom i elev-elev-relationen. Detta resultat implicerar att TH-kombinationen enligt bedömarna skulle vara den mest lämpliga påverkningskombinationen för att åstadkomma beteendeförändringar. För att möjliggöra en granskning av bedömarnas medelbedömning med hänsyn till perception (a_1) och värdering (a_2), när det gäller de enskilda items, utfördes en kontrastanalys för faktorkombinationen AV, som har presenterats i grön bilaga 4, dock utan kommentarer.

Analysen av precisionen och styrkan i de signifikanta F-testen granskades. Resultatet visade, att dessa effekter utgör ett säkrare tolkningsunderlag än lkk självbedömningar. Att experternas bedömning har resulterat i dels större effekter, dels högre sannolikheter är dock ej något oväntat resultat. Detta resultat kan hänföras dels till att bedömarna har en mera homogen referensram, inte minst på grund av den bedömarträning som föregick bedömningen av mikrolektionerna, dels till att variabiliteten i medelbedömningarna av naturliga skäl (k = 2) är mera begränsad, dvs skillnaderna mellan cellmedelvärdena behöver

ej vara av samma storleksordning som för lkk för att resultera i signifikanta effekter.

De signifikanta F-kvoterna utgör sammanfattande test för skillnader i bedömarnas medelbedömning och är dessutom ej riktade. Först en granskning av de signifikanta F-testen genom kontrastanalyser ger information om vad som har förorsakat dessa signifikanser. Denna granskning har resulterat dels i flera medelstora och stora effekter, dels i högre styrkevärden än motsvarande analys av lkk självbedömningar (del 2). Av de skäl som anfördes ovan, har dock kontrastanalyserna, försedda med några kommentarer, redovisats i bilageform för att ge den intresserade läsaren tillfälle att närmare studera cellmedelvärdena. I anslutning till denna redovisning sammanfattas resultatet i grön bilaga 3, tabell 20.

23. Avslutande diskussion av ANOVA-resultaten

Vid analysen av bedömarnas medelbedömning har vi försökt att explicit påvisa, i vilken utsträckning det statistiska resultatutfallet kan anses utgöra ett beviskraftigt tolkningsunderlag. Bedömarnas medelbedömningar har granskats med hjälp av ANOVA-modellen.

23.1 Huvudeffekter

Resultatanalysen visade, att externt förmedlad självkonfrontation via ITV/VR enligt bedömarnas medelbedömning i ingen av de sex olika subjekt-objektrelationerna har lett till iakttagbara skillnader. För faktor T accepteras alltså H_0 för de enskilda variabelområdena.

Orsakerna till ett accepterande av H_0 för samtliga sex variabelområden kan naturligtvis vara många. Det är dessutom mycket svårt att ge en tillfredsställande förklaring till en accepterad nollhypotes.

Den traditionella handledningen i form av en dyadisk konfrontation (faktor H) har däremot lett till påvisbara skillnader i lkk beteenden. Medelbedömningen indikerar skillnader, dels i elev-ego-relationen, dels i elev-elev-relationen. En granskning av cellmedelvärdena har visat, att lkk beteende avviker i positiv riktning, vilket innebär, att det förekommit färre socialt provokativa elevbeteenden, och att elevernas ordningsbeteenden (disciplin) var bättre än i de övriga grupperna. Båda huvudeffekterna visar dock enbart små ES (effektstorlek), och med hänsyn till överskattningen av g-värdena är detta resultat ett relativt osäkert tolkningsunderlag.

1. Tolkas resultatet mot bakgrunden av att handledaren har koncentrerat sin handledning på lkk undervisningsbeteenden avseende elevaktivering, kan effekterna ses som ett resultat av handledningsprocessen.
2. Tolkas resultatet mot bakgrund av att den traditionella handledningen förutsätter att handledaren finns med i klassrummet medan undervisningen pågår, kan man ej bortse från möjligheten, att dessa effekter har uppkommit på grund av att handledarens närvaro har haft en dämpande effekt, dels på provocerande beteenden hos eleverna mot lk, dels på elevernas beteenden sinsemellan.

Tolkningsalternativ 2 förefaller vara det mest troliga, eftersom ingen av de relationer där lk är subjekt har lett till påvisbara effekter.

Mikrolektion 1 och 2 som variationsorsak (faktor U) har enligt medelbedömningen lett till påvisbara effekter avseende ego-IPO-, elev-ego- och elev-elev-relationerna. Som cellmedelvärdena visat, går förändringen dock ej i samma riktning. I ego-IPO-relationen går förändringen i positiv riktning medan den i elevdimensionen (4, 5) går i negativ riktning. Detta resultat implicerar, att alla lkk, oavsett den experimentella påverkan, i samband med lektionstillfälle 2 har visat en iakttagbar förbättring i undervisningsmetodiken, medan eleverna har visat ett mera störande beteende. Precisions- och styrkegranskning visar dock, att resultatet även i detta fall ger ett föga beviskraftigt tolkningsunderlag, möjligen med undantag av elev-ego-relationen, vilket ger vid handen att eleverna vid ml_2 har visat ett iakttagbart mera "provocerande" beteende.

Trots att man ur tolkningssynpunkt måste betrakta effekterna i faktor U som osäkra, kan de vara värda några spekulativa kommentarer. Det måste förefalla något egendomligt, att elevbeteendena blir mera störande, trots att lk undervisningsmetodik blir bättre. För att närmare kunna studera denna utveckling krävs det trendstudier, som dock ej kan utföras inom ramen för denna undersökning. Man skulle nämligen mycket väl kunna anta att, oavsett hur välvilligt undersökningsmetodik än bedöms, så har den föga betydelse för lärar-elev-relationen eller med andra ord för klassrumsklimatet. En tänkbar orsak till detta resultat skulle också kunna vara att första lektionstillfället inneburit en viss anspänning för både lkk och elever. Vid andra lektionstillfället, då lk och eleverna bättre kände varandra, kan en avslappning ha medfört ett mera "vardagsnära" beteende, som har bedömts som ett mera störande beteende.

För aspekterna i mätinstrumentet, dvs bedömarnas perception och värdering (faktor A), kan konstateras, dels mycket stora effekter, dels mycket höga sannolikheter för effekternas storlek (utom för elev-IPO-relationen, där perception och värdering tycks skilja sig obetydligt). Detta resultat är dock i detta sammanhang av mindre betydelse, eftersom man borde kunna förvänta sig, att perceptionen överlag skiljer sig från värderingen. Granskningen av cellmedelvärdena för denna faktor visade, att bedömarnas genomsnittliga perception av beteendeaspekterna inom de enskilda variabelområdena har resulterat i positiva markeringar. Därutöver indikerar bedömarnas medelbedömning att dessa beteendeaspekter är väsentliga i undervisningssammanhang.

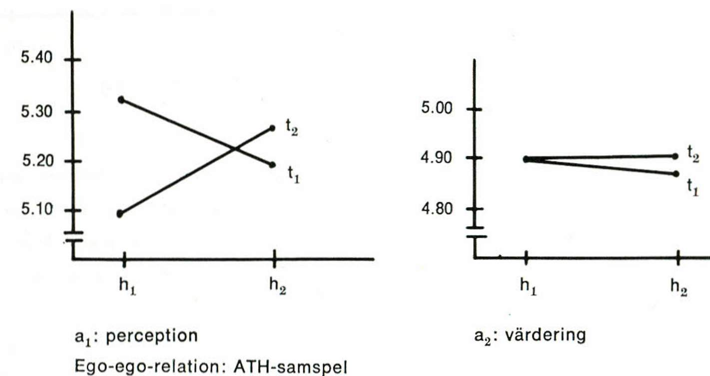
De enskilda variablerna (faktor V) har ej gjorts till föremål för någon mera detaljerad analys, för att undvika att resultattolkningen blir alltför detaljerad.

23.2 Samspelseffekter

Externt förmedlad självkonfrontation via ITV/VR i kombination med traditionell handledning i form av en dyadisk konfrontation, dvs faktorkombination TH, har enligt bedömarnas medelbedömning lett till iakttagbara skillnader i lkk beteenden. Medelbedömningen indikerar, utom för ego-IPO-relationen (variabelområde 3) och elev-elev-relationen (variabelområde 5), skillnader i lkk beteenden till följd av denna faktorkombination (TH).

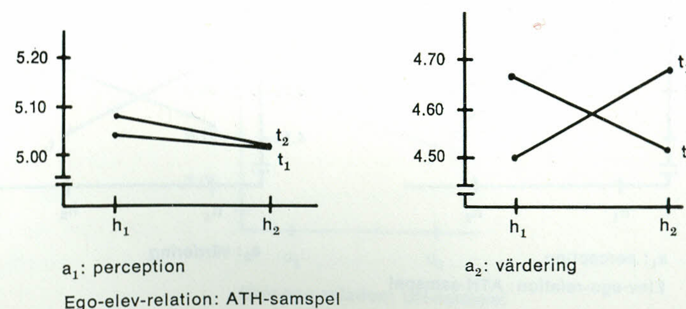
Precisions- och styrkeskattningen implicerar dock, att en eventuell resultattolkning skulle bygga på föga beviskraftiga effekter.

ATH-samspelen visar däremot genomgående medelstora och stora effekter. Styrkan i dessa effekter är dessutom tillräckligt hög för att man skall kunna tolka dem. Bedömningen bör dock relateras till den diskussion som har förts i samband med resultatsammanfattningen eftersom $k=2$. Vid prövningen av de enkla effekterna accepteras för ATH-samspelen i samtliga fall H_0. Nedan åskådliggörs dessa effekter grafiskt. Till varje enskild figur ges en kort kommentar.



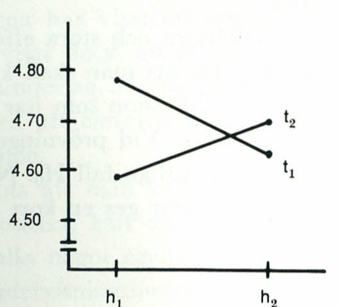
a_1: perception a_2: värdering
Ego-ego-relation: ATH-samspel

Inom ego-ego-relationen implicerar bedömarnas perception (a_1), att t_1h_1-kombinationen åstadkommit den mest positiva effekten. De lkk som ej fick någon handledning tycks visa ungefär samma beteende. Bedömarnas värdering (a_1) skiljer sig däremot knappast med hänsyn till de enskilda experimentgrupperna.

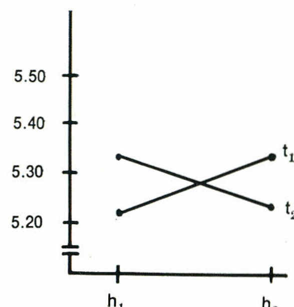


a_1: perception a_2: värdering
Ego-elev-relation: ATH-samspel

Inom ego-elev-relationen tycks ett omvänt förhållande föreligga. Medan skillnaden i perceptionen av lkk beteenden är liten, finns det en relativt kraftig effekt i värderingen (a_2). Samspelet i a_2 skulle kunna förklaras med att experimentet genomfördes i två faser. De mikrolektioner som videobandades under experimentets första fas (år 1969) bedömdes av de pedagogiska experterna samma år, medan de mikrolektioner som videobandades år 1970 också bedömdes år 1970. Samspelseffekten i värderingen (a_2) skulle således kunna hänföras till en förändring i bedömarnas värderingsstruktur till följd av detta årsintervall.



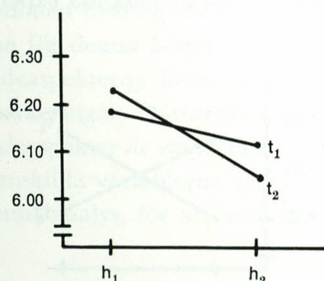
a_1: perception



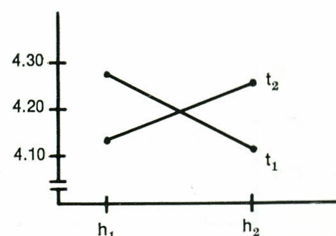
a_2: värdering

Ego-IPO-relation: ATH-samspel

Inom ego-IPO-relationen återspeglar samspelet samma tendens som inom ego-ego-relationen. Bedömarnas perception (a_1) är mest positiv för de lkk som utsattes för externt förmedlad självkonfrontation via ITV/VR i *kombination med* traditionell handledning (t_1h_1). Effekterna i värderingen (a_2) tycks även i detta fall ha uppstått genom att bedömningen har utförts med ett års intervall. Att därmot samspelet visar ett omvänt förhållande mellan perception och värdering är ett intressant resultat, vars orsak dock ej kan utredas på basis av denna resultatanalys. I vilken utsträckning även effekterna i perceptionen måste hänföras till en förändring i bedömarnas perceptionsstruktur kan dock ej avgöras. Denna felkälla är ej under experimentell kontroll.



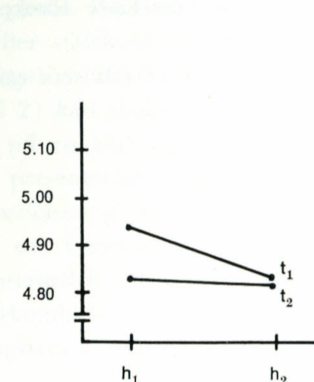
a_1: perception



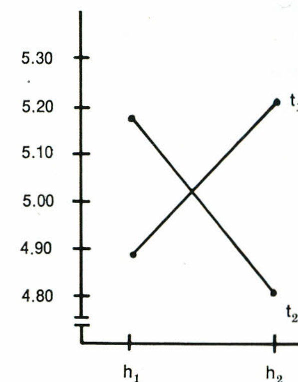
a_2: värdering

Elev-ego-relation: ATH-samspel

Inom elev-ego-relationen implicerar ATH-samspelet med hänsyn till perceptionen (a_1) en tendens till att ett socialt provokativt beteende från eleverna har förekommit i något mindre utsträckning inom de grupper som ej fick handledning än inom de grupper som fick handledning. För värderingen (a_2) gäller däremot samma argumentation som ovan, nämligen att bedömarnas skattningar är förknippade med experimentets två faser.



a_1: perception



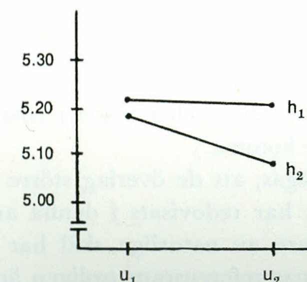
a_2: värdering

Elev-IPO-relation: ATH-samspel

Inom elev-IPO-relationen tycks traditionell handledning *i kombination med* externt förmedlad självkonfrontation via ITV/VR (t_1h_1) enligt bedömarnas perception (a_1) ha lett till ett mera positivt beteende än de övriga påverkningskombinationerna. För värderingen (a_2) gäller samma argumentation som ovan, dvs samspelseffekten torde ha uppstått på grund av intervallen mellan bedömningarna.

Förutom de ovan diskuterade samspelseffekterna ger medelbedömningen uttryck för påvisbara effekter i ytterligare faktorkombinationer, som dock alla hänför sig till elevdimensionen. Precisions- och styrkegranskningen visade dessutom tillfredsställande värden.

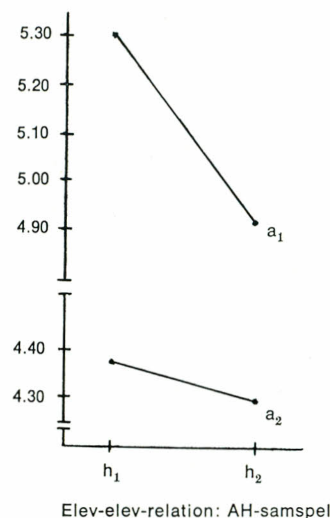
För elev-ego-relationen har ett UH- och ett AH-samspel påvisats. I båda fallen accepteras dock nollhypotesen för de enkla effekterna. UH-samspelet visar följande relationer:



Elev-ego-relation: UH-samspel

Som framgår av den grafiska presentationen implicerar medelvärdena en tendens till en negativare medelbedömning av de lkk som ej fick någon handledning (h_2) vid lektionstillfälle 2. Möjliga förklaringar till denna utveckling presenterades i samband med diskussionen om huvudeffekterna. AH-samspelet är av mindre intresse. Cellmedelvärdena indikerar att det ej i någon nämnvärd utsträckning har förekommit socialt provokativa elevbeteenden, medan sådana skulle bedömas som obesvärande om de hade förekommit. Inom de enskilda aspekterna är skillnaderna ytterst små.

Inom elev-elev-relationen har som enda samspelseffekt påvisats ett AH-samspel, vars relationer är följande:



Elev-elev-relation: AH-samspel

Granskningen av precisionen och styrkan i denna effekt visar dessutom tillfredsställande värden. H_0 för de enkla effekterna har förkastats för skillnaden inom perceptionen (a_1) dvs $a_1h_2 < a_1h_1$. Vi kan alltså konstatera, att det enligt bedömarnas perception har förekommit ett mera disciplinerat elevbeteende när lkk fick handledning. Dels skulle man kunna tolka denna effekt som en framgångsrik handledningsprocess, dels kan man tänka sig att effekten hänför sig till det faktum att handledaren satt med i klassen. (Här bör kanske erinras om, att expertpanelen varken haft kännedom om vilka lkk som hörde till resp påverkningsgrupp eller visste i vilken ordningsföljd de enskilda mikrolektionerna skulle komma.)

Sammanfattningsvis kan sägas, att de överlag större ES-värdena och högre sannolikheterna för ES, som har redovisats i denna analys, måste ses i relation till (1) att två bedömare av naturliga skäl har en lägre spridning än 96 lkk och (2) att bedömarnas referensram troligen är mera enhetlig än lkk, inte minst på grund av bedömarträningen. En felkälla i medelbedömningarna, som ej har varit under experimentell kontroll, är det faktum att expertbedömningarna av tekniska skäl genomfördes med ett års intervall. Som framgått har tydligen denna olägenhet medfört vissa effekter som kan ha påverkat ett mera renodlat resultatutfall.

Slutligen kan tänkas, att denna resultatanalys ger upphov till ett blandat mottagande. De läsare, som av något skäl favoriserar traditionell handledning och/eller självkonfrontation via ITV/VR kan tänkas uppleva resultatet som ett negativt utfall. De läsare, som sätter det i relation till resultatanalys 1 (i del 2) kan tänkas uppleva resultatet som ett väsentligt och informativt bidrag till en kartläggning av lkk utbildningssituation. Ser man nämligen de ovan presenterade resultaten och konklusionerna som en bekräftelse på lkk självbedömning (med bedömarnas mera differentierade referensram i minnet), så är det besynnerligt att ej mera övertygande effekter har kunnat påvisas. Traditionell handledning (H) eller självkonfrontation via ITV/VR (T) samt faktorkombinationen (TH) har ej förmått påverka lkk till att åstadkomma iakttagbara undervisningsbeteenden, som skiljer sig nämnvärt från varandra.

DEL 4

ANALYSER AV LÄRARKANDIDATERS SJÄLVBEDÖMNING OCH PEDAGOGISKA EXPERTERS BEDÖMNINGAR

24. Lärarkandidaters självbedömning och pedagogiska experters medelbedömning

Analysen av lkk självbedömning (del 2) och analysen av pedagogiska experters (bedömare nr 1 och nr 2) bedömningar (del 3) baserades på den grundläggande hypotesen, att experimentets videobandade undervisningssituationer och lkk respektive bedömares observationer strukturellt har ett inbördes förhållande.

I del 2 granskades huvudsakligen om den experimentella påverkan hade lett till påvisbara skillnader i fpp perception och värdering av sina egna, under experimentet utförda undervisningsprestationer. En viktig målsättning i lärarutbildningen är att utveckla lkk färdighet i att tolka undervisningsprocessen "realistiskt", dvs "objektivt". Det antas, att en jämförelse av individens observationer med en individspecifik referensram (en subjektiv måttstock) bestämmer individens aktuella handlande i form av t ex undervisningstaktiker. Skulle man vilja studera vilken information som i realiteten finns tillgänglig genom experimentets videobandade undervisningssituationer eller med andra ord "objektivitetsgraden" i fpp perception och värdering, blir ett yttre kriterium nödvändigt, vilket i denna studie kan betraktas som ett "objektivitetsmått". Detta yttre kriterium utgörs i experimentet av pedagogiska experters medelbedömning.

I del 3 granskades bedömaröverensstämmelsen hos de enskilda pedagogiska experterna, dvs "objektiviteten" i bedömarnas perception och värdering. Dessutom granskades med hjälp av ANOVA om fpp undervisningsbeteenden skiljer sig påvisbart från varandra när bedömarnas medelbedömning utgör jämförelsekriteriet.

Den operationella definitionen av "objektiva" perceptioner och värderingar utgörs således av bedömaröverensstämmelsen mellan å ena sidan enskilda pedagogiska experter och å andra sidan panelens medelbedömning och fpp självbedömning. Hur objektiviteten i fpp observationer påverkas av enskilda faktorer och faktorkombinationer kommer att granskas i föreliggande resultatanalys (3). I de utförda beräkningarna har som objektivitetsmått använts differenserna mellan de pedagogiska experternas medelbedömning och fpp självbedömning.

I resultatanalysens första del redovisas och diskuteras en variansanalytisk granskning av objektivitetsmåttet. ANOVA-modellen är densamma som för de tidigare redovisade analyserna. Det innebär, att också hypoteserna gäller i oförändrad form:

H_0: traditionell handledning ger ingen effekt
H_0: självkonfrontation via ITV/VR ger ingen effekt
H_0: samspel mellan traditionell handledning och självkonfrontation via ITV/VR föreligger ej.

I analysens andra del studeras med hjälp av en kanonisk korrelationsanalys det strukturella sambandet i båda bedömargruppernas observationsdata. Definieras fpp självbedömning genom variabeluppsättning l och bedömarnas medelbedömning genom variabeluppsättning b, kan vi studera den strukturella överensstämmelsegraden mellan l och b. Den kanoniska korrelationsanalysmodellen har kommit till användning för att visa (1) *om båda datauppsättningarna påvisbart är relaterade till varandra och* (2) *på vilket sätt dessa variabeluppsättningar kan kombineras med varandra så att korrelationen mellan komponenterna maximeras.*

Kanoniska korrelationsanalyser är något ovanliga i beteendevetenskapligt sammanhang, varför valet av en sådan analysmodell kanske kräver någon motivering. Är problemställningen, som i denna undersökning, den att två relativt stora variabeluppsättningar skall studeras med hänsyn till variablernas inbördes förhållande, är man framför allt intresserad av några få lineära kombinationer i varje grupp. De variabelkombinationer som är högst korrelerade granskas därvid i första hand. Modellen medför dessutom att strukturen i regel nästan fullständigt kan beskrivas med de första kanoniska variablerna, dvs med ett fåtal okorrelerade lineära kombinationer. Modellen innebär, att relationen mellan båda variabeluppsättningarna har reducerats till sin enklaste form, varför metoden tycks vara speciellt användbar i samband med explorativa studier. För en närmare beskrivning av metodens tillämpning hänvisas till kap 27 och 28.

25. ANOVA av skillnaderna mellan lärarkandidaters självbedömning och pedagogiska experters medelbedömning

För att man skall kunna studera objektiviteten i fpp självbedömning, definierad genom bedömarnas medelbedömning, krävs det att protokollmaterialet och skattningsformuläret (kategorierna) för båda bedömargrupperna är identiska. Har fpp bedömt sin egen undervisning "objektivt", borde det ej finnas någon numerisk skillnad mellan fpp och bedömarnas skattningsvärden. Förekommer skillnader, implicerar dessa, enligt den operationella definitionen ovan, att fpp självbedömning ej överensstämmer med "realiteten". Analys- och redovisningsproceduren är densamma som tillämpades i den första och i den andra resultatanalysen, dvs i del 2 och 3. (Om ej något annat anges, refereras i del 4 till de vita bilagor som har redovisats i Bierschenk 1972 e.)

Mot bakgrund av den tidigare förda diskussionen genomfördes granskningen av ANOVA-tabellerna också i denna analys i följande tre steg

1. tolkning av mönstret i F-testen
2. granskning av precisionen och styrkan i F-testen
3. efterprövningar.

25.1 Mönstret i F-testen

Resultatet av de för de enskilda variabelområdena separat utförda variansanalyserna sammanfattas i tabell 68. F-kvoterna granskades i syfte att beskriva tolkbara mönster. F-test ($\alpha = .05$) redovisas i detta sammanhang dock enbart för att ett mönster skall kunna framträda tydligare. Som framgår ur tabellen, har ANOVA för de enskilda variabelområdena resulterat i ett antal signifikanta skillnader med hänsyn till objektiviteten, så som den definierats i denna analys. För att visa om en eventuell struktur i de signifikanta F-kvoterna tillåter en meningsfull tolkning granskades tabellen både radvis och kolumnvis.

En radvis utförd granskning av huvudeffekterna visar, att det ej finns någon påvisbar skillnad i objektiviteten, som kan hänföras till antingen (1) externt

Tabell 68. Skillnaderna mellan lärarkandidaternas självbedömning och bedömarnas medelbedömning: Sammanfattning av ANOVA-tabellerna

Variations-orsak	Variabelområde 1	2	3	4	5	6
T						
H						
TH	*			*		
U	**			**		
UT						
UH					**	*
UTH						
A		**		**		
AT	*		**			
AH						
ATH	*	**		**	**	**
AU			*	*		
AUT			**			
AUH						**
AUTH						
V	**	**	**	**	**	**
TV						
HV						
THV	*	*			**	
UV						
UTV						
UHV						
UTHV						
AV	**	**	**	**	**	*
ATV						
AHV						**
ATHV						*
AUV			*			
AUTV						
AUHV						
AUTHV						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
V: Skattnings- och värderingsformulär F III
* $F_{.99}(1, 92) = 7.08$
** $F_{.95}(1, 92) = 4.00$

förmedlad självkonfrontation via ITV/VR (faktor T) eller (2) traditionell handledning i form av dyadisk konfrontation (faktor H). Med utgångspunkt från de hypoteser som ställdes ovan accepteras nollhypotesen för båda faktorerna, vilket betyder att objektiviteten i fpp perception och värdering av sin egen undervisning ej påverkas av vare sig externt förmedlad självkonfrontation via ITV/VR eller traditionell handledning i form av en dyadisk konfrontation.

Följer man faktor T i de enskilda variabelområdena, finner man ett AT-samspel inom ego-ego-relationen (1) och inom ego-IPO-relationen (3). Dessutom utläses inom ego-IPO-relationen ett signifikant AUT-samspel.

Följer man faktor H i de enskilda variabelområdena i fråga om sådana faktorkombinationer där faktor H enbart förekommer med faktor A och/eller U, visar sig följande resultat: Ett UH-samspel inom elev-elev-relationen (5) och elev-IPO-relationen (6) har resulterat i påvisbara effekter. Därutöver finner man en signifikant effekt i AUH-samspelet inom variabelområde 6.

Mönstret i F-kvoterna visar, att de påvisbara skillnaderna i fpp objektivitet på ett regelbundet sätt finns antingen inom *ego*-dimensionen eller inom *elev*-dimensionen. Samspelseffekterna med faktor T som variationsorsak i variabelområdena 1—3 implicerar, att självkonfrontationsprocesserna har medfört, att lkk objektivitet varierar på ett systematiskt sätt inom de relationer, i vilka *ego* är subjekt. Samspelseffekterna med faktor H som variationsorsak implicerar, att lkk som utsattes för dyadisk konfrontation (traditionell handledning) på ett systematiskt sätt påverkas i sin objektivitet vid bedömningen av den egna undervisningen, men avvikelserna hänför sig nu till de relationer där *elev(er)* är subjekt (variabelområdena 4—6). Någon närmare tolkning är dock i detta analyssteg ej möjlig utöver konstaterandet att sådana effekter finns.

För TH-samspelet finner man i tabell 68 två påvisbara effekter ($\alpha = .05$), inom ego-ego-relationen (1) och inom elev-ego-relationen (4). Dessa skillnader implicerar, att självkonfrontation via ITV/VR *och* traditionell handledning har påverkat objektiviteten. Men eftersom TH-samspelet i kombination med precisionsfaktor (A) inom ego-ego-relationen inte heller resulterade i någon större effekt, tycks avvikelserna i perception och värdering vara av mindre betydelse. De övriga ATH-effekterna implicerar däremot avvikelser som uppfyller kriteriet ($\alpha = .01$). Anmärkningsvärt är, att denna effekt har uppstått inom samtliga variabelområden, där eleverna är subjekt i relationen. Inom ego-dimensionen finner man däremot en effekt endast inom ego-elev-relationen (2), vilket anger att objektiviteten har påverkats när fpp bedömer eleverna.

Undervisningstillfällena, dvs mikrolektion 1 och 2, anges i analysmodellen med faktor U. Denna faktor har resulterat i två signifikanta effekter, nämligen inom ego-ego-relationen (1) och elev-ego-relationen (4). Dessa avvikelser implicerar, att fpp objektivitet, oavsett om fpp har fått externt förmedlad självkonfrontation eller traditionell handledning, har förändrats från mikrolektion

1 till mikrolektion 2. Anmärkningsvärt är också, att båda dessa effekter rör de variabelområden, där ego är objekt. En närmare granskning av huvudeffekterna och vissa samspelseffekter redovisas i kapitel 26.

25.2 Precisionen och styrkan i F-testen

Resultatet, som framgår ur tabell 68 ovan indikerar, att det finns signifikanta skillnader i objektiviteten med avseende på fpp perception och värdering. De signifikanta effekterna implicerar dock ej att associationsgraden mellan de experimentella faktorerna och mätningen måste vara av lika storleksordning för samma signifikansnivå. Ett mått på precisionsgraden i dessa mätningar är PV, som kan skattas med hjälp av Hays $\hat{\omega}^2$. PV, som hänför sig till de enskilda effekterna (tab 68), redovisas i tabell 69.

Tabell 69. $\hat{\omega}^2$-värden för signifikanta F-kvoter: summavariabel

Variations-	Variabelområde					
orsak	1	2	3	4	5	6
T						
H						
TH	(.02)			(.02)		
U	.02			.01		
UT						
UH					.02	(.01)
UTH						
A		.12		.04		
AT	(.01)		.01			
AH						
ATH	(.01)	.04		.03	.02	.06
AU			(.01)	(.00)		
AUT			.02			
AUH						.02
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
(): Anger $F_{.95}$ (1, 92) = 4.00

Som framgår ur tabell 69 har endast faktor A inom variabelområde 2 (ego-elev) ett $\hat{\omega}^2$-värde, som avviker markant från de övriga effekterna. Även prediktorvariansen i faktorkombinationen ATH för variabelområde 6 (elev-

Tabell 70. Effektstorlek och sannolikhet för skillnaderna mellan lärarkandidaternas och bedömarnas skattningar

Variations-	Variabelområde					
orsak	1	2	3	4	5	6
1. Effektstorlek (f)						
T						
H						
TH	(.16)			(.16)		
U	.18			.15		
UT						
UH					.26	(.16)
UTH						
A		.37		.16		
AT	(.11)		.19			
AH						
ATH	(.24)	.43		.28	.24	.48
AU			(.16)	(.16)		
AUT			.39			
AUH						.38
AUTH						
2. Sannolikhet (g)						
T						
H						
TH	(.35)			(.35)		
U	.78			.61		
UT						
UH					.84	(.35)
UTH						
A		>.99		.67		
AT	(.15)		.52			
AH						
ATH	(.40)	.93		.55	.40	.93
AU			(.35)	(.35)		
AUT			.88			
AUH						.85
AUTH						

T: Externt förmedlad självkonfrontation via ITV/VR-tekniken
H: Traditionell handledning (dyadisk konfrontation)
U: Undervisningstillfälle (mikrolektion 1, 2)
A: Aspekt (perception, värdering)
( ) $F_{.95}$ (1, 92) = 4.00

IPO) ger uttryck för en relativt stor del av den systematiska variansen. Mellan de övriga oberoende variablerna och den beroende variabeln är PV däremot av samma storleksordning som för de resultat som har redovisats i del 2 och 3. $\hat{\omega}^2$ presenterades som ett proportionsmått, vars numeriska värde är en funktion av antalet effekter i analysen. Om t.ex. en faktor med repeterad mätning sätts ur spel blir ω^2 för de resterande effekterna större än de ω^2 beräknas utan sammanslagning.

Effektstorleken för en viss effekt, när alla övriga effekter hålles konstanta, kan skattas med hjälp av Cohens f (Cohen, J., 1969, ss 266—276). Beräknas f, blir det också möjligt att med hjälp av Cohens tabellverk skatta sannolikheten, dvs styrkan i F-testen. En sammanfattning av storleksordningen för respektive signifikant F-test, ES (f) och styrkeskattningen (g) redovisas i tabell 70.

Som framgår ur tabell 70 är ES i de flesta variationsorsakerna medelstor och sannolikheten i ett antal fall tillfredsställande ($g > .70$). Är styrkan för en viss effekt lägre, måste ett accepterande av respektive alternativ hypotes betraktas som tvivelaktigt.

25.3 Efterprövningar

Analyssteg 1 och 2 utgör ett viktigt hjälpmedel vid tolkningen av signifikanta F-test. Först sådana granskningar möjliggör en bedömning av huruvida analysresultaten kan anses vara värda ett mera detaljerat studium av vad som förorsakat resultatutfallet. Analyssteg 1 har visat ett meningsfullt mönster i F-testen, och analyssteg 2 visar både ett antal medelstora effekter och tillfredsställande styrkevärden. En mera utförlig resultatanalys och -diskussion är alltså med hänsyn till vissa effekter ($g > .70$) befogad. Kontrastanalyserna, dvs tredje steget i analysen, har utförts med hjälp av Newman—Keuls och Scheffés metod. Men även för sådana effekter, som ej har uppfyllt kriteriet för att man skall kunna betrakta dem som ett beviskraftigt tolkningsunderlag, har efterprövningar utförts. Alla efterprövningar redovisas, försedda med några kommentarer, i vit bilaga 3. Att även kontrastanalyserna för de ur tolkningssynpunkt tvivelaktiga effekterna redovisas i form av en bilaga betyder, att vi dels ville följa upp de signifikanta F-testen, som är baserade på $F_{.99}\,(1, 92) = 7.08$, dels ville (för dem som så önskar) bereda tillfälle till en detaljerad granskning av samspelseffekterna.

26. Sammanfattning av ANOVA-resultaten

Vill man studera "objektiviteten" i fpp självbedömning, krävs det något "objektivitetsmått". Den operationella definitionen av "objektiva" perceptioner och värderingar avsåg i denna studie pedagogiska experters medelbedömning. Har fpp bedömt sin egen undervisning "objektivt", borde det enligt den ovan presenterade definitionen ej finnas någon numerisk skillnad mellan fpp och bedömarnas observationsvärden. I denna del av resultatanalysen redovisades sex ANOVA av skillnaderna mellan fpp självbedömning och pedagogiska experters (bedömare nr 1 och nr 2) medelbedömning. I anslutning till såväl den första som den andra resultatanalysen (del 2 och 3) redovisades en sammanfattning av resultaten, och en avslutande diskussion fördes. I resultatredovisningen ovan har därför diskussionen enbart berört skillnaderna med avseende på objektiviteten. För diskussion av sådana problem som antingen gäller lkk självbedömning eller bedömarnas medelbedömning hänvisas till de tidigare presenterade resultatanalyserna.

Bedömarna och fpp har utfört sina skattningar enligt skattnings- och värderingsformulär F III. Granskningen av ANOVA-tabellerna skedde i enlighet med den tidigare presenterade analystekniken, i tre steg. Mönstret i F-testen visade, att objektiviteten i fpp självbedömning i ett antal faktorer och faktorkombinationer påvisbart varierar. Inom de enskilda subjekt-objektrelationerna har dock varken faktor T eller faktor H medfört någon påvisbar skillnad. Variationer i fpp objektivitet vid bedömningen av den egna undervisningen har alltså i detta avseende ej kunnat påvisas.

I faktor U fanns däremot påvisbara effekter inom variabelområdena 1 och 4, vilket innebär att fpp objektivitet signifikant förändras från mikrolektion till mikrolektion.

I faktor A fanns signifikanta effekter inom variabelområdena 2 och 4, vilket innebär att fpp i sin perception och värdering, oavsett den experimentella påverkan och lektionstillfällena, ej är "objektiva".

Precisions- och styrkegranskningen visade dels små effekter (i ett fall medelstor), dels måttliga styrkevärden med hänsyn till att g-värdena för huvudeffekterna utgör en viss överskattning av sannolikheten för effektstorleken (ES). Enbart effekter som återspeglar en skillnad i användningen av mät-

instrumentets aspekter inom ego-elev-relationen kan således betraktas som ett säkert underlag för tolkningen.

De signifikanta samspelseffekterna i faktorkombinationen TH ($\alpha = .05$) har varit av intresse endast i fråga om mönstret i F-testen. Men trots att ATH-samspelen visade medelstora och stora ES är ur tolkningssynpunkt endast effekterna inom ego-elev-relationen och inom elev-IPO-relationen ($g > .70$) tillräckligt beviskraftiga. Dessutom kan AUT-samspelet inom ego-IPO-relationen, UH-samspelet inom elev-elev-relationen och AUH-samspelet inom elev-IPO-relationen betraktas som ett beviskraftigt tolkningsunderlag. Effekterna i dessa faktorkombinationer är medelstora. Ett intressant resultat, som mönstret i F-testen återspeglat, är, att ES i de sistnämnda faktorkombinationerna i samband med T hänför sig till ego-dimensionen, medan kombinationerna med H hänför sig till elev-dimensionen.

Eftersom F-kvoter utgör sammanfattande test, som dessutom ej är riktade, kan man först genom kontrastanalyser få en mera ingående information om vad som har förorsakat dessa signifikanser. En sammanfattning av kontrastanalyserna ges i tabell 17 (se vit bilaga 3:13).

27. Avslutande diskussion av ANOVA-resultaten

Den variansanalytiska granskningen av objektiviteten i fpp självbedömning har utförts för att visa eventuella variationer i de differensvärden som beräknats på basis av fpp självbedömning och bedömarnas medelbedömning. Oavsett om det inom var och en av dessa datauppsättningar har funnits påvisbara effekter eller ej, skulle man nämligen mycket väl kunna tänka sig att den experimentella påverkan har medfört skillnader, dvs variationer i fpp "objektivitet". I samband med denna resultatanalys har vi, såsom tidigare, försökt att explicit påvisa i vilken utsträckning det statistiska resultatutfallet kan betraktas som ett säkert tolkningsunderlag.

27.1 Huvudeffekter

Denna resultatanalys har ej visat någon signifikant skillnad i fråga om de experimentella faktorerna T och H. För båda faktorerna accepteras alltså inom samtliga subjekt-objekt-relationer H_0, dvs faktorerna har ej påverkat objektiviteten i fpp självbedömning. I bedömningen av i vilken utsträckning det egna undervisningsbeteendet har förändrats från lektionstillfälle 1 till 2 (faktor U) varierar dock fpp bedömning påvisbart i förhållande till objektivitetskriteriet. Cellmedelvärdena visar, att skillnaden är störst i samband med mikrolektion 1 och att den minskar relativt kraftigt vid bedömningen av mikrolektion 2. För ego-ego-relationen saknas tyvärr, på grund av ett utskrivningsfel från datorn, tillgång till bedömarnas cellmedelvärden. Inom elev-ego-relationen visar cellmedelvärdena följande förändring: Vid lektionstillfälle 1 är fpp bedömning mera negativ än bedömarnas, medan fpp bedömning i samband med lektionstillfälle 2 blir mera positiv och bedömarnas mera negativ. Även om detta resultat ej får läggas till grund för någon tolkning, visar sig dock en intressant tendens: medan fpp upplever en förbättring, sker enligt bedömarna en försämring. Här bör erinras om att bedömningspanelen ej haft någon kännedom om vilken lektion som var fpp första respektive andra.

Visserligen implicerar minskningen av skillnaden i samband med mikrolektion 2 en mera enhetlig och kanske också mera objektiv bedömning. Ur pedagogisk-psykologisk synvinkel och med hänsyn till lärarutbildningen bör man dock lägga märke till att lkk och pedagogiska experter närmar sig kongruensen i bedömningen genom att självbedömningen blir mera positiv medan medelbedömningen blir mera negativ. Ur lk utgångsläge tycks det vara själv-

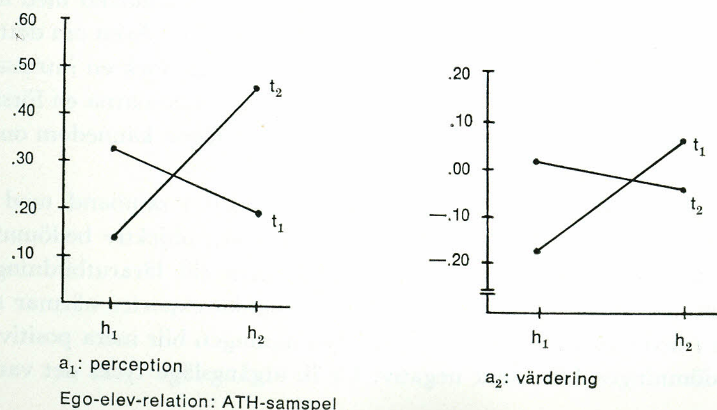
klart, att det har skett en framgång vid undervisningstillfälle 2. Enligt pedagogiska experter är så dock inte alls fallet.

Av de signifikanta effekterna i faktor A utgör enbart effekten inom ego-elev-relationen ett säkert tolkningsunderlag. Skillnaden betyder att fpp har varit mera negativa än bedömarna i fråga om perceptionen (a_1), medan värderingen i stort sett har varit densamma.

Trots att effekten inom elev-ego-relationen ur tolkningssynpunkt är mindre beviskraftig, kan den vara värd en kommentar. Elev-ego-relationen definieras huvudsakligen av sådana items, som beskriver ett socialt provokativt elevbeteende. Både fpp och bedömarna är i sin perception tämligen överens om att ett sådant beteende knappast har förekommit. I värderingen av i vad mån ett sådant beteende skulle vara besvärande för undervisningen skiljer sig dock fpp påvisbart från bedömarna. Medan fpp värderar ett socialt provokativt beteende som ganska besvärande, värderas det av bedömarna som relativt obesvärande. Kommer den fortsatta forskningsverksamheten att bekräfta hypotesen att individens ”själv” utgör ett effektivt kontroll- och regleringsinstrument för en kontinuerlig förbättring av undervisningen, får ett sådant resultat ökad betydelse. Eftersom läraren genom sin ledarfunktion på ett avgörande sätt påverkar den socialpsykologiska strukturen, måste nödvändigtvis hans interpretation av vad som är ett socialt provokativt beteende samt hans värdering av sådant beteende få ökad uppmärksamhet i utbildningssammanhang.

27.2 Samspelseffekter

Som framgått av resultatsammanfattningen uppfyller enbart vissa av de signifikanta samspelseffekterna kraven på att ES skall kunna betraktas som ett säkert tolkningsunderlag. Denna diskussion avser således enbart effekter vars styrka är $g > .70$. För ego-dimensionen uppfyller inom ego-elev-relationen ATH-samspelet och inom ego-IPO-relationen AUT-samspelet detta krav. ATH-samspelets utseende kan åskådliggöras på följande sätt:



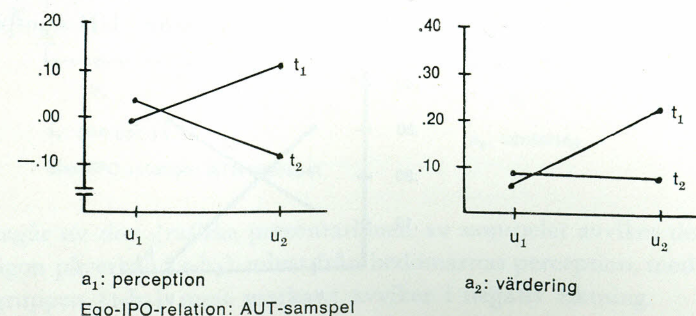
Ego-elev-relation: ATH-samspel

Kontrastanalysen för ATH-samspelet har ej resulterat i någon påvisbar skillnad mellan cellvärdena, varför H_0 accepteras. Avseende perceptionen indikerar cellmedelvärdena en tendens till att de fpp som ej fick någon påverkan alls (t_2h_2) i större utsträckning avviker från bedömarnas perception än övriga fpp. Minst skiljer sig de fpp som fick påverkan i form av traditionell handledning (t_2h_1) från bedömarnas perception. Denna tendens skulle kunna implicera, att någon form av påverkan (traditionell handledning i större utsträckning än övriga påverkningsformer) leder till en mera ”objektiv” perception, definierad genom en given pedagogisk-psykologisk norm.

Som den grafiska presentationen ovan visar, värderas i den aktuella mikrolektionen väsentlighetsgraden inom dessa beteendeaspekter på följande sätt: De fpp som fick påverkan i form av traditionell handledning (t_2h_1) avviker från experternas värdering obetydligt i positiv riktning. De fpp som fick påverkan i form av externt förmedlad självkonfrontation via ITV/VR (t_1h_2) avviker däremot i något större utsträckning i positiv riktning. Ingen påverkan alls (t_2h_2) resulterade i att fpp värderade beteendeaspekterna mera negativt än experterna. Den jämförelsevis största avvikelsen, som dessutom går i negativ riktning, hänför sig till faktorkombinationen t_1h_1.

Som den grafiska presentationen visar, existerar ett omvänt förhållande mellan perception och värdering. Skulle man vilja tolka tendensen i detta samspel, implicerar resultatet, att värderingsstrukturen påverkas av handledaren så, att den bäst överensstämmer med den aktuella pedagogiska normen, medan en kombination av externt förmedlad självkonfrontation via ITV/VR och traditionell handledning (TH) tydligen inom ego-elev-relationen har medfört den mest markanta avvikelsen inom a_2. Men eftersom skillnaderna mellan cellmedelvärdena ej är signifikanta, bör man ej fästa större uppmärksamhet vid detta resultat.

Kontrastanalysen för AUT-samspelet inom ego-IPO-relationen har inte heller resulterat i påvisbara skillnader mellan cellmedelvärdena. AUT-samspelet kan åskådliggöras på följande sätt:



Ego-IPO-relation: AUT-samspel

Den tendens som framträder i perceptionen och i värderingen implicerar, att såväl de fpp som fick påverkan i form av externt förmedlad självkonfrontation via ITV/VR (t_1) som de fpp som ej fick sådan påverkan (t_2) vid lektionstillfälle 1 ganska väl överensstämmer med experterna (bäst överensstämmer t_1-gruppen), medan denna överensstämmelse minskar för lektionstillfälle 2. De fpp som tillhör t_1-gruppen percipierade sig själva mera positivt än experterna gjorde, medan t_2-gruppen vid lektionstillfälle 2 i större utsträckning avviker i negativ riktning.

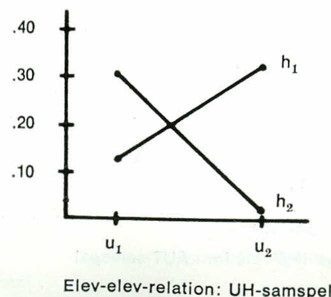
Kanske skulle man kunna tolka resultatet som en tendens som bestyrker hypotesen att fpp förväntar sig en förbättring av undervisningsmetodiken vid lektionstillfälle 2 och att denna förväntning uppleves som bestyrkt via ITV/VR, oavsett det "objektivt" rådande förhållandet, medan de fpp som tillhör t_2-gruppen ej har samma möjlighet att genom TV-monitorn kunna "se" sina förväntningar bestyrkta.

Även för värderingsaspekten visar kontrastanalysen en intressant tendens fastän H_0 för cellmedelvärdena accepteras. t_1- och t_2-gruppen avviker vid lektionstillfälle 1 i positiv riktning från experterna, och denna avvikelse är lika stor för t_2-gruppen vid mikrolektion 2. t_1-gruppen värderar däremot vid mikrolektion 2 beteendeaspekterna markant mera positivt än experterna. Värderingen tycks i detta fall följa perceptionen, dvs i den mån man percipierar en framgång med hänsyn till undervisningsmetodiska aspekter, värderas dessa också som mera väsentliga. Den omvända relationen gäller dock ej, eftersom man, trots att perceptionen är mera negativ, vidhåller värderingsnivån.

Tolkningen baseras visserligen på icke signifikanta cellmedelvärden, men tendensen tycks peka i samma riktning som forskningsresultat som har presenterats i annat sammanhang. Diggorys (1966, s 229) experimentella arbeten visade, att en framgångsrik lösning av en uppgift medför en mera frekvent "självvärdering" uppåt, medan ett misslyckande mindre frekvent medför motsvarande värdering nedåt.

För elev-dimensionen uppfyller inom elev-elev-relationen UH-samspelet och inom elev-IPO-relationen ATH- och AUH-samspelen kraven för att effekterna kan betraktas som tillfredsställande tolkningsunderlag.

Som framgått ur kontrastanalysen för UH-samspelet accepteras H_0, dvs cellmedelvärdena skiljer sig ej påvisbart från varandra. UH-samspelet kan åskådliggöras på följande sätt:

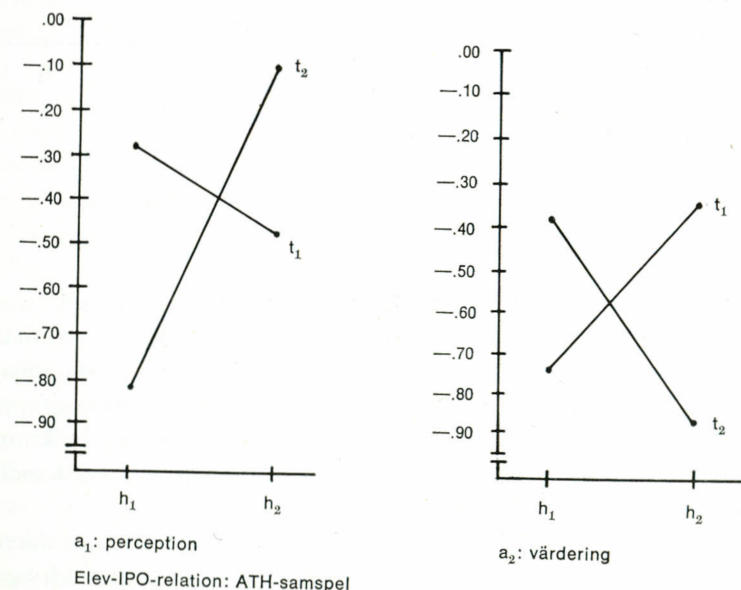


Elev-elev-relation: UH-samspel

Tendensen i cellmedelvärdena implicerar, att de fpp som ej fick påverkan i form av traditionell handledning (h_2) vid lektionstillfälle 1 i större utsträckning avviker från experternas bedömning än de fpp som fick påverkan i form av traditionell handledning (h_1). Vid lektionstillfälle 2 implicerar medelvärdena en intressant förändring: h_1-gruppens bedömning avviker kraftigt från bedömarnas, medan h_2-gruppen knappast visar någon avvikelse alls. I samtliga fall avviker fpp från bedömarnas bedömning i positiv riktning.

Att de fpp som tillhör h_1-gruppen avviker från bedömarna skulle kunna förklaras med att handledarens kommentarer och bedömning av undervisningen ej överensstämmer med bedömarnas och att fpp dessutom delvis har tolkat in handledarens synsätt i den via ITV/VR förmedlade undervisningssituationen. En möjlig förklaring till förändringen i motsatt riktning för h_2-gruppen skulle kunna vara, att TV-monitorn har medfört en normerande effekt. En mindre spekulativ förklaring än denna kräver dock dels ytterligare analyser, som t ex granskning av den slumpmässiga tilldelningen av faktor T resp faktor H i fält 1, dels trendstudier med flera än två mikrolektioner.

Kontrastanalysen av ATH-samspelet inom elev-IPO-relationen visar, att fpp i sin perception och värdering avviker från bedömarnas i negativ riktning. Någon signifikant skillnad mellan grupperna kan dock ej påvisas. ATH-samspelet har följande utseende:

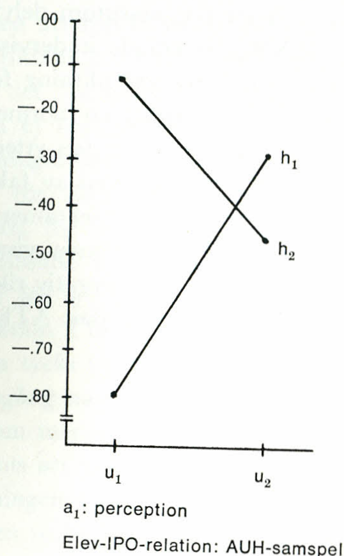


Elev-IPO-relation: ATH-samspel

Som framgår av den grafiska presentationen av samspelet avviker de fpp som ej fick någon påverkan (t_2h_2) minst från bedömarnas perception, medan handledningsgruppen (t_2h_1) mest markant avviker i negativ riktning.

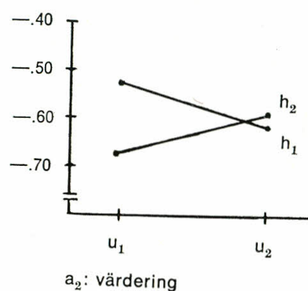
I värderingen återspeglar cellmedelvärdena ett omvänt förhållande, eftersom de fpp som ej fick någon påverkan (t_2h_2) i sin värdering i högre grad avviker från bedömarnas. Även för t_1-gruppen kan det konstateras en omvänd relation. Jämföres detta resultat med ATH-samspelet inom ego-elev-relationen, visar sig samma omvända relation, fastän skillnaden inom elev-IPO-relationen såväl för a_1 som för a_2 är mera utpräglad.

Slutligen finns det inom elev-IPO-relationen ett AUH-samspel som kan åskådliggöras på följande sätt:



a_1: perception

Elev-IPO-relation: AUH-samspel



a_2: värdering

Kontrastanalysen visar en signifikant cellmedelvärdesskillnad, nämligen $u_1h_2 < u_1h_1$. Denna skillnad i perception implicerar, att de fpp som fick handledning (h_1) vid lektionstillfälle 1 (u_1) varit påvisbart mera negativa än de fpp som ej fick handledning (h_2). Medan det för handledningsgruppen vid lektionstillfälle 2 kan konstateras en relativt stor förändring i riktning mot experternas perception, har skillnaden för de fpp som ej fick handledning ökat i negativ riktning.

Någon påvisbar skillnad i cellmedelvärdena för värdering finns däremot ej. Cellmedelvärdena implicerar därutöver mycket obetydliga förändringar från mikrolektion 1 till mikrolektion 2.

För h_1-gruppens perception visar sig tydligt vid lektionstillfälle 2 en förändring i riktning mot bedömarnas perception. För värderingsaspekten kan däremot konstateras en ökning av skillnaden. Som framgår av AUH-samspelet återspeglar cellmedelvärdena för h_2-gruppen (ej handledning) en omvänd tendens. Jämfört med pedagogiska experter ökar för dessa fpp skillnaden i perceptionen, medan den minskar i värderingen.

Sammanfattningsvis kan sägas, att tendensen i de samspelseffekter, som har granskats närmare, implicerar, att handledning inom ego-dimensionen medförde en ökad överensstämmelse mellan fpp och de pedagogiska experternas perception och en ökad skillnad i värderingen. Jämför man externt förmedlad självkonfrontation via ITV/VR med övrig påverkan, har såväl för perceptionen som för värderingen en ökning i skillnaden vid mikrolektion 2 kunnat konstateras.

Inom elev-dimensionen tycks skillnaden mellan h_1-gruppen och h_2-gruppen öka inom elev-elev-relationen vid lektionstillfälle 2, medan skillnaden minskar för elev-IPO-relationen, fastän handledningsgruppen närmar sig bedömarna, medan avståndet i bedömningen för h_2-gruppen ökar.

Som resultatanalyserna 1, 2 och 3 implicerar, finns det till följd av den experimentella påverkan endast små mätbara effekter. Förklaringarna till dessa resultat skulle kunna vara många. Kerbers (1967, s 32) undersökning ”*Study of selected student teachers and modifications of their self-concept through use of portable video tape recordings*” visade samma resultat. Han skriver, att det fanns påtagligt små mätbara förändringar i fpp ”self-concept” till följd av en externt förmedlad konfrontation via ITV/VR. Enligt författaren tycks den största förändringen kunna påvisas i det som fpp betraktar som egenskaper hos en *ideal*lärare. Att studera möjliga förändringar i fpp lärarideal har dock ej varit planerat, varken i Kerbers undersökning eller i det aktuella experimentet. Att handledaren ej har lyckats påverka fpp i högre grad skulle kunna bero på att fpp ännu ej har utvecklat lämpliga testkriterier, dvs pedagogisk-psykologiska normer. Men det kan naturligtvis också förklaras med ett fenomen som Stoller (1970, s 11) observerade i terapeutiskt sammanhang. Enligt honom undviker både handledare och kandidat en kritisk analys genom att de använder sig av ord som ”a verbal portrait of an individual” i syfte att undvika att kritiskt granska det egna beteendet, bl a genom att det under den dyadiska konfrontationen diskuteras händelser, som ligger utanför den egentliga gruppsituationen, t ex undervisningsprocessen. Att den externa självkonfrontationen ej omedelbart medför förändringar antingen i personens normsystem eller aktuella beteende kan möjligen förklaras med att det fanns ”a trained incapacity to deal with first-hand, non-verbally mediated self-information” (Stoller, 1970, s 11).

För att information skall kunna förmedlas, måste det alltså finnas en strukturell interrelation mellan interna och externa värden. Med denna hypotes som utgångspunkt har protokollmaterialet granskats i resultatanalyserna 1—3. Perlmutter m fl (1967, ss 900—905) rapporterar, att individer, involverade i en gruppsituation, som medlemmar är begränsade i sin perception genom in-

kongruensen mellan perceptionen av det som händer i själva grupprocessen och det som observeras via ITV/VR. Denna situation medför, enligt författarna, att personerna ifråga är benägna att förändra i första hand perceptions- och värderingsstrukturen (normsystemet), medan de aktuella beteendena bibehålls. Men allteftersom individen bygger upp ett nytt självförtroende, kommer inkongruensen mellan individens normsystem och det observerade beteendet att minskas genom att individen försöker förändra sitt eget beteende, t ex genom att bli mera självkritisk och lättare acceptera kritiska kommentarer. Tendensen i analysresultaten indikerar, att det i första hand finns tecken på förändringar i perceptions- och värderingsstrukturen. Men bedömarnas observationer indikerar också beteendeförändringar. Orsaken till att det ej finns mera entydiga resultat är troligen primärt en följd av den korta påverkningstiden. Dels har lkk enbart haft tillfälle till 2 mikrolektioner, dels har hela experimentet genomförts inom en tidsrymd av 6 veckor. Som ett experiment, rapporterat av Stoller (1970, s 179), visar, är det mycket möjligt att beteendeförändringar förekommer (enligt externa bedömare), trots att fpp självbedömning, såsom den återspeglas i testresultaten, ej indikerar någon form av förändring. I en första fas av en konfrontationsprocess behöver dessutom eventuella modifikationer av individens perceptions- och värderingsstruktur ej uppträda samtidigt med eventuella modifikationer av individens beteendemönster.

28. Kanonisk korrelationsanalys av lärarkandidaters självbedömning och pedagogiska experters medelbedömning

För att kunna lära sig olika alternativa handlingar måste individen först successivt lära sig bygga strukturer. Detta måste ske genom en sammankoppling av interna strukturer med externa. I denna process används den interna strukturen som kriterium för att testa externa strukturer, såsom beteenden eller handlingsmönster, som har realiserats. Kriteriet antas bestå av individens perceptions- och värderingsstruktur, medan den externa strukturen t ex skulle kunna bestå av realiserade beteenden eller handlingsmönster i undervisningssituationer.

Som framgått av de hittills redovisade analyserna har denna jämförelseprocess blivit manifesterad i form av olika observationsdata. Skattningsvärden föreligger dels för fpp, dels för pedagogiska experters perception och värdering av experimentets videobandade protokollmaterial.

Medan ANOVA har utförts i syfte att visa olika individers selektion av information till följd av den experimentella påverkan, avses med hjälp av kanoniska korrelationsanalyser ett studium av möjliga strukturella likheter mellan pedagogiska experters och lkk observationsdata.

För att man skall kunna predicera en icke mätbar variabel (t ex skolframgång) med hjälp av en grupp observerbara variabler (t ex betyg), presenterades av Hotelling (1935, ss 139—142) den kanoniska korrelationsanalysmodellen. Trots att denna analysmodell blev känd redan 1935, har den mycket sällan kommit till användning inom den beteendevetenskapliga forskningen. Metoden har beskrivits bl a av Bartlett, 1941, ss 29—38; Thomson, 1947, ss 27—34; Anderson, 1958, ss 288—306; Kendall, 1965, ss 75—85; Cooley & Lohnes, 1971, ss 168—192. För diskussionen av metoden hänvisas till dessa beskrivningar.

Problemställningen i samband med de kanoniska korrelationsanalyserna är att studera relationen mellan lkk självbedömning och pedagogiska experters medelbedömning. Analysförfarandet kan man utföra i åtminstone tre olika steg. *Första* steget innebär, att man granskar om det över huvud taget finns bivariata relationer, som är signifikanta. Finns sådana, utförs analysens andra

steg. *Andra* steget innebär, att man granskar på vilket sätt de olika variablerna har bidragit till den signifikanta relationen. Variablerna kan således studeras (1) dels med hänsyn till om det finns lika höga koefficienter eller vikter för en viss variabel i båda datauppsättningarna, dvs om strukturen överlappar perfekt, (2) dels på vilket sätt en eller flera delgrupper i bedömarnas variabeluppsättning (b) är relaterade till en eller flera delgrupper i lkk variabeluppsättning (l). Svaret på frågan om det finns speciella kombinationer av variabler eller variabelgrupper i b, som är relaterade till variablerna eller variabelgrupper i l, bestäms genom en granskning av de enskilda variablerna som ingår i datauppsättningen. Medan man med hjälp av Bartletts χ^2-test (Cooley & Lohnes, 1971, s 175) kan bestämma hur många funktioner som tillåter en statistisk tolkning, är gränsdragningen mera godtycklig i fråga om variablernas vikter. Finns det olika variabler eller variabelgrupper som är relaterade till varandra, utförs tredje steget i analysen. *Tredje* steget innebär ett försök att ge dessa relationer ett meningsfullt innehåll. Men här uppstår samma svårigheter som man möter vid tolkningen av komponentanalyser, faktoranalyser etc, nämligen problemet att avgöra, huruvida dessa lineära funktioner motsvarar någon "verklighet" eller om de måste ses som ett matematiskt resultat ("artefacts") som saknar innehåll. Vid tolkningsförsöket hade det varit till stor hjälp om sådana analyser hade tillämpats i empirisk forskning i större utsträckning än vad som hittills har skett.

Innan de olika kanoniska korrelationsanalyserna utfördes, granskades de enskilda variablernas medelvärden, spridningar och skalor. Medelvärdena och spridningar redovisas för hela datauppsättningen i Bierschenk, 1972 e, vit bilaga 5. De variabler som eventuellt skulle ha kunnat invalidera analysen och av detta skäl har uteslutits, redovisas för perception (a_1) i tabell 71 och för värdering (a_2) i tabell 72.

Tabell 71. Items som har uteslutits ur kanoniska korrelationsanalyserna, perception (a_1)

Item nr	m	s	max	min
	Bedömarnas medelbedömning, ml_1			
11	7.00	.00	7.00	7.00
28	6.65	.44	7.00	4.50
30	1.72	.42	3.00	1.00
33	3.92	.40	5.00	2.50
34	6.78	.44	7.00	4.50
61	6.83	.48	7.00	4.00
62	6.99	.05	7.00	6.50
63	6.97	.15	7.00	6.00
64	6.94	.37	7.00	4.00
66	6.98	.11	7.00	6.00
67	6.98	.11	7.00	6.00
70	6.78	.40	7.00	5.00
	Lärarkandidaternas självbedömning, ml_1			
62	6.86	.49	7.00	3.00
	Bedömarnas medelbedömning, ml_2			
11	6.99	.05	7.00	6.50
28	6.61	.38	7.00	5.50
30	1.80	.44	4.00	1.00
33	3.93	.37	5.00	2.50
34	6.81	.37	7.00	5.50
61	6.95	.19	7.00	5.50
62	6.98	.09	7.00	6.50
63	6.95	.19	7.00	6.00
64	7.00	.00	7.00	7.00
66	6.95	.24	7.00	5.50
67	6.96	.17	7.00	6.00

ml : mikrolektion
m : medelvärde
s : standardavvikelse

Som kriterium för uteslutningen bestämdes att $s < .50$ eller variationsvidd < 1.00. Detta kriterium valdes med utgångspunkt från experimentets råmaterial.

Tabell 72. Items som har uteslutits ur kanoniska korrelationsanalyserna, värdering (a_2)

Item nr	m	s	max	min
	Bedömarnas medelbedömning, ml_1			
19	6.44	.21	6.50	5.00
34	2.93	.24	4.00	2.50
36	6.26	.27	6.50	5.50
50	5.58	.26	6.50	5.00
57	6.24	.25	6.50	6.00
66	4.03	.26	5.50	3.00
67	3.98	.19	4.50	3.00
	Bedömarnas medelbedömning, ml_2			
8	3.90	.28	4.50	2.50
11	5.20	.28	5.50	4.00
19	6.43	.19	6.50	5.50
37	5.96	.17	6.50	5.00
57	6.25	.27	6.50	5.50
61	4.52	.16	5.50	4.00
66	4.02	.23	4.50	3.00
67	3.99	.23	4.50	3.00

ml : mikrolektion
m : medelvärde
s : standardavvikelse

Som framgår av Bierschenk, 1972 e, vit bilaga 5 är spridningen för värderingen (a_2) genomgående lägre än för perceptionen (a_1), varför i denna analys beslutet för uteslutningen baserades på kriteriet $s < .30$ eller variationsvidd < 1.00. Detta kriterium valdes med utgångspunkt från experimentets råmaterial.

Efter att de i tabellerna 71 och 72 redovisade items hade rensats bort, utfördes för varje enskilt variabelområde, undervisningstillfälle och aspekt en kanonisk korrelationsanalys. För varje analys testades med $\alpha = .05$ om det fanns signifikanta korrelationer. Resultatet redovisas i tabellerna 73 och 74.

Tabell 73. Signifikansprövning för kanoniska korrelationer (R_c), perception (a_1), mikrolektion 1 och 2

Variabel-område	R_c	Observerat χ^2-värde	Kritiskt χ^2-värde	$\alpha = .05$
Mikrolektion 1				
Ego-ego	1	429.35	406.00	*
	2	351.15	366.17	
Ego-elev	1	574.28	490.69	*
	2	405.57	447.36	
Ego-IPO	1	247.65	147.94	*
	2	125.51	124.06	*
	3	90.99	103.29	
Elev-ego	1	28.41	26.01	*
	2	12.55	16.63	
Elev-elev	1	33.44	26.01	*
	2	6.18	16.63	
Elev-IPO	1	18.93	26.01	
Mikrolektion 2				
Ego-ego	1	437.63	406.00	*
	2	364.62	366.17	
Ego-elev	1	459.91	490.69	
Ego-IPO	1	141.78	147.39	
Elev-ego	1	30.63	26.01	*
	2	9.16	16.63	
Elev-elev	1	40.53	26.01	*
	2	13.50	16.63	
Elev-IPO	1	19.51	26.01	

Som framgår ur tabell 73 finns det för perceptionen (a_1), ml_1 i ego-dimensionen för varje variabelområde en signifikant kanonisk korrelation. Inom variabelområde 3 finns det dessutom en andra signifikant korrelation. I elev-dimensionen har, utom för variabelområde 6, den första korrelationen blivit signifikant. I samband med lektionstillfälle 2 har korrelationerna förändrats. I ego-dimensionen finns det enbart en signifikant korrelation (ego-ego-relation). I elev-dimensionen är den första korrelationen inom variabelområdena 4 och 5 signifikant.

Tabell 74. Signifikansprövning för kanoniska korrelationer (R_c), värdering (a_2), mikrolektion 1 och 2

Variabel-område	R_c	Observerat χ^2-värde	Kritiskt χ^2-värde	$\alpha = .05$
Mikrolektion 1				
Ego-ego	1	415.29	406.00	*
	2	357.04	366.17	
Ego-elev	1	424.08	447.36	
Ego-IPO	1	131.41	147.39	
Elev-ego	1	10.26	26.01	
Elev-elev	1	16.48	26.01	
Elev-IPO	1	13.49	26.01	
Mikrolektion 2				
Ego-ego	1	378.13	406.55	
Ego-elev	1	490.69	453.37	*
	2	326.51	447.36	
Ego-IPO	1	117.25	147.39	
Elev-ego	1	10.63	26.01	
Elev-elev	1	20.95	26.01	
Elev-IPO	1	26.91	26.01	*
	2	6.34	16.63	

En granskning av tabell 74 visar, att det finns ett mycket mindre antal signifikanta korrelationer för värderingen (a_2). För lektionstillfälle 1 har enbart den första dimensionen i ego-ego-relationen lett till ett signifikant resultat. Vid lektionstillfälle 2 sker en förändring såtillvida, att det nu i stället i ego-dimensionen finns en signifikant relation inom variabelområde 2. Dessutom kan även för elev-dimensionen påvisas en korrelerad dimension, eftersom den första kanoniska korrelationen inom variabelområde 6 har blivit signifikant.

I de fall där den kanoniska korrelationen har blivit signifikant, kan en påvisbar interrelation konstateras mellan variabeluppsättning b och l med $\alpha = .05$. I början av tillämpningen av kanoniska korrelationsanalyser trodde man att enbart den första signifikanta dimensionen och respektive kanonisk korrelation skulle vara av intresse. Men utvecklingen av dessa första tillämpningar visar, att även andra dimensioner än den första kan vara relevanta för tolkningen. Resultatet av de kanoniska korrelationsanalyserna visar, att det finns sammanlagt 12 signifikanta korrelationer, dvs 12 dimensoner i bedömarnas medelbedömning, som signifikant är relaterade till respektive dimensioner i fpp självbedömning.

De kanoniska korrelationer, som har blivit signifikanta granskades med hänsyn till vikterna inom respektive variabeluppsättning. Vid granskning av överensstämmelsen med hänsyn till enskilda items, kan man tillämpa olika strikta krav. Ett kriterium skulle kunna vara, att vikterna parvis är av samma

storleksordning eller att de överensstämmer inom en viss toleransgräns. För att man skall kunna betrakta överensstämmelsen som relativt exakt bestämdes toleransgränsen till $\pm.10$.

Skalorna i formulär F III är bipolära. Ett lågt värde på respektive skala motsvaras av det första adjektivet, och ett högt värde motsvaras av den pol som definieras av det andra adjektivet i adjektivparet. En hög positiv vikt för en viss skala anger således, att en fp som ligger närmast den andra polen i respektive dimension troligen kommer att få höga skattningsvärden på den kanoniska variabeln ifråga. En hög negativ vikt implicerar, att en fp ligger närmast polen, som definieras av det första adjektivet och troligen kommer att få ett lågt skattningsvärde på denna dimension.

Empiriska data leder sällan till entydiga resultat. Detta faktum minskar dock ej den kanoniska korrelationsanalysens potentiella värde. Blandade resultat för tecknens riktning betyder, att dimensionerna i den ena variabeluppsättningen (b) som är högt korrelerad med den andra variabeluppsättningen (l) ej nödvändigtvis medger en meningsfull verbal beskrivning.

28.1 Kanoniska korrelationer och koefficienter för de enskilda subjekt-objekt-relationerna

I vit bilaga 5, tabellerna 1—9 (Bierschenk, 1972 e) redovisades de signifikanta kanoniska korrelationerna och koefficienterna för respektive subjekt-objekt-relation. Korrelationsmatriserna redovisades i vit bilaga 7:1—7:2, tabellerna 1—24. Som framgår ur tabellerna i bilaga 5:1—5:8 har koefficienterna för de enskilda variablerna såväl mellan viktvektorerna som inom respektive vektor resulterat i olika förtecken. Denna teckenblandning medför, att vektorerna ej kan betraktas som uttryck för något allmänt karakteristiskt som är definierat genom de enskilda komponenterna. Hade vikterna mellan de enskilda komponenterna resulterat i vikter med gemensamma, positiva förtecken, så skulle man ha kunnat kombinera komponenterna till ett för bedömare och lkk gemensamt perceptions- respektive värderingsindex inom variabelområdet ifråga. De erhållna vikterna tillåter dock ej sådana kombinationer. Nästa steg i analysen innebär en granskning av strukturen i respektive kanonisk variabel. Som framgår ur tabellerna är vikterna i de enskilda variablerna dels positiva, dels negativa, vilket medför att man inte heller kan konstruera separata perceptions- respektive värderingsindex, dvs ett bedömare- respektive ett lärarkandidatsindex.

Ett undantag från den ovan beskrivna situationen finns dock inom ego-IPO-relationen. Inom detta variabelområde har den kanoniska korrelationsanalysen resulterat i två signifikanta dimensioner. Korrelationerna och vikterna för respektive dimension framgår ur tabell 4 (vit bilaga 5:4). Som tabellen visar, hänför sig den första korrelationen tydligen helt och hållet till item nr

55, nämligen användningen av tavlan. Att använda tavlan viktas såväl i b_1- som i l_1-variabeln positivt. Överensstämmelsen mellan båda variabeluppsättningarna är perfekt. Dessutom kan den första dimensionen förklaras enbart med detta item, eftersom alla andra vikter kan anses vara utan betydelse.

Den första kanoniska korrelationen i ego-IPO-relationen illustrerar hurudant resultatutfallet borde ha varit för att man skulle ha kunnat konstruera meningsfulla index. Ett försök att kombinera vikterna i den andra signifikanta dimensionen till ett meningsfullt index möter däremot samma svårigheter som har anförts ovan.

Sammanfattningsvis kan sägas, att resultatet av den kanoniska korrelationsanalysen visar, att det finns signifikant korrelerade dimensioner i pedagogiska experters och lkk perception och värdering. Dessa har sedan granskats, för att sådana påståenden i respektive kanonisk variabel skall kunna utkristalliseras, som skulle kunna ingå i perceptions- respektive värderingsindex. Genom dels "för stora" skillnader mellan de parvisa vikterna, dels olika förtecken har de enskilda påståendena dock ej kunnat kombineras till tolkbara index. Detta resultat kan mycket väl tänkas bero på, att regressionsvikterna bestäms så, att korrelationen blir maximal.

Koefficienter för de signifikanta korrelationerna har redovisats i bilaga 5 i syfte att för den intresserade läsaren möjliggöra en granskning av variablernas relativa position inom de enskilda vektorerna.

Mot bakgrund av detta resultat skulle man kunna fråga sig om vissa restriktioner med hänsyn till viktvektorerna skulle kunna medföra resultat som bättre motsvarar målsättningen att etablera en gemensam struktur.

Såsom den kanoniska korrelationsanalysen är definierad, söker modellen efter *den optimala lösningen* vid maximeringen av korrelationerna, vilket kan resultera i "artefacts". Ett tänkbart alternativt tillvägagångssätt skulle kunna vara att man inför restriktioner i modellen. Man skulle t ex kunna beräkna enbart en viktvektor, dvs för varje enskilt item skulle vikten vara densamma för bedömare och fpp. På detta sätt skulle man kunna erhålla korrelationer mellan två items enligt hypotesen att vikterna är lika. Den kanoniska korrelationsanalysen skulle på så sätt bli av hypotesprövande natur och leda till en suboptimal lösning. Man kan då ställa frågan hur mycket lägre den kanoniska korrelationen blir och om en något lägre korrelation har praktiska konsekvenser eller ej. Det är därmed inte sagt att tolkningen av vikterna nödvändigtvis blir lättare, men ett sådant förfaringssätt skulle genom sin *hypotesprövande natur* onekligen vara mera ändamålsenligt, eftersom det ger upplysningar om korrelationens storlek när två variabeluppsättningar får samma viktning. Ur forskningsmetodisk synpunkt vore det önskvärt, att få tillfälle att närmare undersöka en sådan modifiering av den kanoniska korrelationsanalysmodellen. Utvecklingen av modellen skulle vara ett värdefullt bidrag till den pedagogisk-psykologiska metodforskningen.

29. Sammanfattning av de kanoniska korrelationsanalysernas resultat

Med hjälp av kanoniska korrelationer och koefficienter granskades överensstämmelsegraden mellan pedagogiska experters och lkk perceptions- och värderingsstruktur. Kanoniska korrelationsanalyser kan resultera i flera korrelerade dimensioner. Tolkningen av dessa behöver man ej nödvändigtvis begränsa till sådana som har visat sig vara signifikanta på den valda signifikansnivån. I den aktuella analysen har enbart tagits hänsyn till korrelationer som blev signifikanta med $\alpha = .05$. Eftersom huvudsyftet var att kartlägga, om det eventuellt fanns gemensamma strukturer i fpp och bedömarnas observationsdata, har på samma sätt som vid granskningen av F-testen för eventuella mönster den lägre signifikansnivån bedömts som mera lämplig. Granskningen kan sammanfattas i följande punkter:

1. Som framgått av resultatredovisningen finns det 12 signifikant korrelerade dimensioner, varav 9 inom perceptionsstrukturen och 3 inom värderingsstrukturen. Tolkningen av resultatet implicerar, att likheterna är större i perceptionsstrukturen än i värderingsstrukturen. Dessutom framgår ur tabell 73 att perceptionsstrukturen med hänsyn till de enskilda variabelområdena visar färre dimensioner med ett inbördes förhållande vid mikrolektion 2. För värderingsstrukturen kan konstateras en ökning med en signifikant korrelerad dimension vid lektionstillfälle 2, men de signifikant korrelerade dimensionerna avser vid mikrolektion 1 respektive mikrolektion 2 olika subjekt-objekt-relationer (variabelområden).
2. Granskningen av de signifikanta korrelationerna utfördes i syfte att konstruera ett för fpp och bedömare gemensamt perceptions- respektive värderingsindex. Som kriterium bestämdes att vikterna skulle överensstämma inom toleransgränsen $\pm .10$. Vikterna i de enskilda dimensionerna, som blev signifikanta, uppfyller dock ej, med något undantag, överensstämmelsekriteriet. Detta resultat tillåter ej en konstruktion av gemensamma index som är tolkbara.
3. Vikterna inom varje enskild dimension har resulterat i olika förtecken, varför inte heller en tolkbar konstruktion av separata bedömar- respektive lärarkandidatindex varit möjlig. Ett exempel på hur resultatutfallet borde

ha varit för att man skulle ha kunnat konstruera gemensamma index, är den första dimensionen inom ego-IPO-relationen. Det för vikterna negativa resultatutfallet bör dock ses mot bakgrund av att denna korrelationsanalys bestämmer vikterna så, att de kanoniska korrelationerna maximeras på ett visst sätt. En utveckling av den kanoniska korrelationsanalysmodellen till en *hypotesprövande modell,* där man enbart bestämmer en viktvektor (parvisa komponenter får samma viktning) och därefter beräknar korrelationerna, skulle kunna vara ett alternativt tillvägagångssätt, mera lämpat för problemställningen ifråga.

30. Avslutande diskussion av de kanoniska korrelationsanalysernas resultat

Resultatet av de kanoniska korrelationsanalyser som utförts av bedömarnas och lkk observationsdata visar, att det i bedömarnas och lkk perceptions- och värderingsstruktur finns signifikant korrelerade dimensioner. Dessa har sedan granskats med hänsyn till påståenden i respektive variabeluppsättning som skulle kunna kombineras. Analysen har dock i viktningen av de enskilda påståendena visat skillnader, som är större än vad som kan accepteras för konstruktion av psykologiskt meningsfulla index.

Perception betyder enligt en mycket allmänt hållen definition en interaktion eller transaktion mellan individ och omgivning. Individen får information från den externa världen, som på något sätt modifierar individens erfarenheter och beteenden (Warr & Knapper, 1968, s 2). Även om perceptionsobjekt numera kan hållas konstanta med hjälp av ITV/VR-tekniken, behöver vi veta mera om vilka kriterier olika personer använder vid perception och värdering av objekt.

Begreppet värdering avser specifika kriterier, som måste vara uppfyllda, för att en beteendestrategi skall kunna utföras eller fullföljas. Man undviker på så sätt "valens"-begreppens begränsning, nämligen att ett objekt alltid måste vara av ett visst värde för individen eller tillfredsställa ett visst behov. Som i kap 7:1 anförts, är det allmänt vedertaget, att externa observatörer kan åstadkomma större objektivitet vid perception och värdering av en viss persons beteenden än vad personen ifråga själv kan. Med utgångspunkt från hypotesen att denna inkongruens i objektiviteten kanske i huvudsak beror på att individen ifråga och de externa bedömarna använder sig av föga överlappande perceptions- och värderingsdimensioner, utfördes för varje enskilt variabelområde en kanonisk korrelationsanalys. Som framgått ur resultatsammanfattningen intar pedagogiska experter och lkk knappast samma relativa position i fråga om värderingsstrukturen. Något positivare är däremot resultatutfallet för perceptionsstrukturen, men analyserna visar att antalet signifikant korrelerade dimensioner från lektionstillfälle 1 till 2 minskar.

Det klassiska tillvägagångssättet att granska den struktur som finns i bedömarnas och fpp observationsdata, skulle ha varit att konstruera ett index (ett vägt medelvärde) och på basis av detta granska relationen mellan båda varia-

beluppsättningarna. Men vikterna för indexkonstruktionen hade man måst välja på arbiträr basis och den resulterande korrelationen skulle då självfallet ha berott på vikterna. Man hade naturligtvis också kunnat använda sig av komponentanalysmodellen och därefter granska de transformerade variablernas inbördes förhållande. Jämfört med komponentanalysmodellen har den kanoniska korrelationsanalysmodellen dock påtagliga fördelar. I stället för att operera med korrelerade data (den intresserade läsaren hänvisas till korrelationsmatriserna i Bierschenk, 1972 e, vit bilaga 7) och efter transformationen försöka utreda effekterna, försöker man med hjälp av den kanoniska korrelationsanalysen utreda effekterna före analysens början genom en transformering av de observerade variablerna till nya variabler, som så långt som möjligt är oberoende av varandra. Vikterna i modellen är determinerade genom kravet att korrelationerna skall maximeras. På så sätt får man indexvärden som ger maximal korrelation mellan de karakteristika som kännetecknar bedömarnas och lkk perceptions- och värderingsstruktur. De empiriska variabeluppsättningarna har alltså fortfarande såsom i komponentanalysmodellen transformerats till nya variabler men ej så, att variansen maximeras på ett speciellt sätt. I den kanoniska korrelationsanalysmodellen blir i stället kovariansen mellan vissa speciella variabler tillhörande båda variabeluppsättningarna maximerad, medan kovariansen för de resterande variablerna reduceras till noll. Modellen innebär med andra ord, att relationen mellan båda variabeluppsättningarna har reducerats till sin *enklaste form*. Metoden tycks vara speciellt användbar i samband med explorativa studier (Anderson, 1958, s 288). Är problemställningen, som i denna undersökning, den att två relativt stora variabeluppsättningar skall studeras med hänsyn till variablernas inbördes förhållande, är man som nämnts i första hand intresserad av några få lineära kombinationer i varje grupp. De variabelkombinationer som är högst korrelerade granskas därvid i första hand. Modellen medför dessutom, att strukturen i regel nästan fullständigt kan beskrivas med de första kanoniska variablerna, dvs med ett fåtal okorrelerade lineära kombinationer.

I kap 1.5 anfördes, att man i den moderna systemforskningen alltmera har tagit intryck av den grundläggande principen att alla informationsprocesser är urvalsprocesser, som står i relation till varandra. Att analysera individens sätt att bl a välja ut och kategorisera tillgänglig information har stått i centrum för den kanoniska korrelationsanalysen. Som framgått finns det i de pedagogiska experternas och lkk observationer signifikant korrelerade dimensioner. Men viktningen av de enskilda påståendena i de kanoniska variablerna indikerar att lkk och de pedagogiska experterna i sina bedömningar knappast intar samma relativa position. Tolkningsresultaten implicerar, att den information som ligger till grund för den dyadiska konfrontationen torde leda till att handledare och lk löper risk att tala förbi varandra i diskussionen om enskilda komponenter som utgör en mera komplex undervisningsprocess.

DEL 5

AN ENGLISH DIGEST: SELF-CONFRONTATION VIA CLOSED-CIRCUIT TELEVISION IN TEACHER TRAINING TOGETHER WITH RECOMMENDATIONS FOR FURTHER RESEARCH

31. Problems

One of the goals of Swedish teacher training in pedagogics according to the directives of the Swedish National Board of Education (SÖ, Klasslärarutbildningen, studieplaner, 1968, s. 393) is

"to combine with other aspects of the training in furthering the personal development and self-knowledge of the aspirant teachers and providing them with a vocational training which, given their individual qualities, will equip them for undertaking the responsibility of educating and teaching the grades for which they are being trained".

It has long been the policy within teacher training to let the student teachers discover for themselves, during their period of supervised teaching practice, the way in which their own behavior influences the teaching process. A well-known phenomenon within the psychology of perception is that the individual's structure of perception and evaluation leads to interpretations of objects and situations that are specific for that individual. One of the aims of the traditional tutorial system is to help the student teachers find out about themselves and about the relationships that exist between the student teacher and his pupils. But the difficulties experienced by both the tutor and the student teacher in recalling "*exactly*" what happened in the practice teaching situation obstruct the fulfillment of this aim.

In recent years, externally mediated self-confrontation by means of closed-circuit television and video-recording (CCTV/VR) has become a popular technique of confrontation. A large number of reports and articles, extremely varied in quality, have been published in many different journals. In addition, a number of mimeographed doctoral theses are available. Special bibliographies, two in English and one in German, on "Television as a technical aid in education and in educational and psychological research" have been published by Bierschenk (1969, 1971 b, 1971 c). A third report including all three, with an introduction in Swedish, appeared in Bierschenk (1971 d). A survey of literature on educational and psychological research into the techniques of audiovisual self-confrontation is given in Chapter 3.

The reactions of the teacher when the desired teaching behavior has been specified and accepted by him have been described and analysed in numerous studies. In contrast, the aim of the present study is to investigate the teacher's reactions when *no* specific norms have been externally and explicitly predetermined.

Like the actual teaching situations, the behavior of teachers and students can and has been studied in many different ways. Moreover, it has become increasingly obvious during the past few years that self-knowledge and self-understanding require quite different research methods than the study of curricula, material-method systems and teaching tasks.

Those working in educational-psychological research have long lacked the means of placing the complex process of teaching under experimental control. The technique of micro-lessons (a technique of reduction) has proved to be a very useful research method. The whole structure of micro-lessons can be manipulated in such a way that different problems can be answered by means of experimental designs. In addition, the CCTV system and the video-recorder make possible new approaches in research methodology for the study of interaction processes.

The main purpose of this study, "Self-confrontation via closed-circuit television", has been to study the reactions of student teachers placed under various experimental conditions when confronted with their own teaching performances, which have been registered by means of closed-circuit television and video-recording. An additional aim has been to study the "degree of objectivity" of the perceptions and evaluations of the student teachers, by examining their selection of information from the video-recorded teaching situations. The studies of effect concerned different forms of feedback, such as dyadic confrontations in the form of traditional tutorship and externally mediated self-confrontation processes via CCTV/VR.

To summarize, the goals have been to:

1. study systematically and under controlled conditions the way in which the student teacher perceives and evaluates the behavior of himself and his pupils in the context of micro-lessons.
   As a result of the treatment of this problem, the individual's "self-concept" and "life-space cognition" have come to occupy a central position in the study.
2. examine systematically the dyadic confrontation process in a tutoring situation.
   Since tutoring plays such a central role in teacher training, a detailed analysis of the pattern of "face-to-face" communication ought to be of considerable importance for educational research.

3. gain experience in the use of closed-circuit television, video-recorders and micro-lesson techniques as aids in research and as teaching methods.
   The advantages of CCTV/VR and micro-lessons have been pointed out in various contexts and need not be further emphasized here. One disadvantage of CCTV, video-recording and micro-lesson techniques is that editing at certain stages can make the protocol material liable to subjectivity. Lighting and sound-recording can also cause difficulties.

For a more detailed discussion, see Chapter 2.

32. The experimental design of the study

Studies dealing with possible ways of using CCTV/VR techniques in educational psychology often seem to be characterized either by faulty designs or by inadequate models for the treatment of data.

Stickell's (1963, p. 46) investigation showed that of 250 data comparisons concerning "televised and face-to-face instruction", only 6 % were based on control group designs with "satisfactory control groups", while 50 % were not based on any experimental control. Stickell's examination showed (p. 48) that only 10 out of 250 comparisons led to interpretable results. Controlled experiments are, however, the "only way of verifying educational improvements" (Campbell & Stanley, 1963, p. 172).

The main problem in the present study has been:

What are the effects of traditional tutoring in the form of dyadic confrontation and/or externally mediated self-confrontation via CCTV/VR on the self-assessment (perception and evaluation) of student teachers?

Those participating in this study were 96 female student teachers, admitted in the autumn terms of 1967 and 1968 to the School of Education in Malmö for training as teachers in grades 4—6. They took part in the experiment at the beginning of their second term.

In order to achieve the maximum degree of control over possible interpretations of the results of the study, a factorial design was drawn up. The different factors of the design are:

Factor H: Traditional tutoring, where
h_1: tutoring
h_2: no tutoring

Factor T: Self-confrontation mediated externally via CCTV/VR, where
t_1: self-confrontation
t_2: no self-confrontation

Factor U: Micro-lessons (length 15 min.), where
u_1: micro-lesson 1
u_2: micro-lesson 2

In order to increase the precision of the design, two additional factors were included in the original design, i.e.

Precision factor V: Assessment and evaluation schedule F III
$v_1, \ldots, v_{79}$ statements of which the measuring instrument consists.

Precision factor A: Aspects of the instrument. Each statement contains two different aspects, where
a_1: perception aspect
a_2: evaluation aspect.

The whole ANOVA model can be written as A, U, T, H, I (TH), V, where I denotes the factor representing the individuals. A summary of the experiment's analysis of variance design is presented in Table 75. In addition, a brief description is given below of factors H, T, and U (for a description of factors I, A, and V, see Chapters 4 and 8.2).

Table 75. The analysis of variance design of the experiment

Index	A	U	T	H	I	V
No. of levels	2	2	2	2	24	79
Size of population	2	2	2	2	∞	79

32.1 Description of factors H, T, and U

Factor H: Traditional tutoring in the form of dyadic confrontation was arranged so as to be similar to the tutoring student teachers receive during their teaching practice. Acting in the same way as during normal teaching practice, the tutor observed the student teachers during the experiment, i.e. made the notes considered necessary for the subsequent tutoring session.

The tutor was allowed the same length of time for discussing the lesson with the subject as is normal in ordinary teaching practice.

Factor T: Externally mediated self-confrontation via CCTV/VR here implies confrontation with one's own behavior in teaching situations which are registered via closed-circuit television and video-recording. The process involves external confrontation with one's own expressive behavior. These experiences of confrontation could be described as a de-automatization of the usual way of seeing one's self. This Factor T can thus be said to involve an external self-distancing in space and time.

In order to avoid or balance possible sources of error associated with technical problems such as camera angle or different methods of editing (spatial selection, temporal selection), two cameras were used, linked via a mixer, for registration of the behavior of the experimental subjects, while the behavior of the pupils was registered via a third static camera. To make it possible to examine the facial expressions of the subject, a close-up was registered by zooming in every third minute. The close-up lasted for 10 seconds. This measure was introduced as a result of a preliminary experiment (cf. Ch. 2) where the student teachers expressed a desire to see themselves in close-up. A more detailed description of the techniques used in the experiment for both recording and playback, together with the arrangement of the apparatus, are given in Chapter 6.

Factor U: Micro-lessons involve three different components, namely (1) pupils, (2) theme of the lesson to be taught and (3) length of lesson.

1. *Pupils* participating in micro-lessons should be representative for the level that the aspirant teacher is going to teach. The pupils (half-classes) that took part in 1969 and 1970 all came from the fourth grade of Munkhätteskolan in Malmö and were divided between the four experimental groups without regard to ability or social class.
2. *The theme of the lesson to be taught* was taken from the subject area of biology. Within this area the animals of northern Sweden were chosen: lemming (I), bear (II), wolf (III), lynx (IV), reindeer (V), golden eagle (VI), grouse (VII) and wolverine (VIII). (The Roman figures given in brackets refer to Table 8.) These subjects were chosen on the recommendation of lecturers in Methodology as being relatively simple. At this level there is no noticeable variation in the technique required either for teaching about the different species or for planning the teaching. An extra advantage in choosing biology was that all the student teachers had access to abundant and very similar concrete illustrative material.
3. *Length of lesson,* i.e. the video-recorded teaching time, was limited to 15 minutes. The student teachers were also allowed an additional 5 minutes warming-up time to get acquainted with the pupils. The short duration (15 min.) of the lesson forced the student teachers to keep to the task at hand, and imposed a natural restriction on too wide a variety of teaching activities.

Table 8 gives a summary of the experimental plan for the first micro-lesson. The second micro-lesson repeats exactly the same plan. A survey of the experimental design is presented in Figure 3.

The factorial design shown in Figure 3 is a more complex form of Campbell's and Stanley's design no. 6, "The Posttest-Only Control Group Design".

This design checks all the eight sources of error (see Campbell & Stanley, 1963, p. 178) that could invalidate the internal validity of the experiment. In addition, the experiment must have external validity if we are to make generalizations on the basis of the result for the population of student teachers in question (cf. Fig. 1).

A detailed discussion of the internal and external validity of the experiment, together with a description and discussion of some empirical results that illuminate certain aspects of the ecological validity of the experiment, has been presented in Chapter 11.

32.2 Data from attitude questionnaires

In order to find out to what extent the experiment was felt to be something exceptional compared to the usual teaching situation, three attitude questionnaires were constructed and administered to (1) the teaching staff at the School of Education, (2) the student teachers in the second term of their training as teachers in grades 4—6 who did not participate in the experiment, and (3) the student teachers in the second term of their training as teachers in grades 4—6 who did participate in the experiment. The answers to the separate questionnaires can be summarized as follows:

Teaching staff at the School of Education in Malmö

The teacher-trainers consider that:

1. the experiment described above is "*very important*" for teacher training
2. the participation of student teachers in such experiments is "*very important*" for teacher training and
3. the School of Education should in the future "*to a very great extent*" carry on research into the use of closed-circuit television in teacher training.

Despite the fact that the experiment was integrated into the teacher training schedule, some members of staff have complained that the experiment interfered with the normal course of their work.

Student teachers who did not participate in the experiment

The answers received from the student teachers who did not participate in the experiment can be summarized in the following way:

1. This group has a positive attitude to the CCTV experimental activity in both 1969 and 1970.

2. The student teachers wish to see their own lessons, registered via CCTV/VR, rather often—a positive attitude to the medium and the method concerned.
3. The student teachers are very hesitant as to whether they could accept losing some of their scheduled training. This reaction is completely in line with the current practice at the School of Education. Furthermore, the student teachers are hardly willing to accept a greater load of work.
4. The risk that the student teachers participating in the first phase of the experiment should have lost essential parts of the training was judged in 1969 to be minimal, while the student teachers participating in the second phase were thought to have lost essential lectures.

   A possible explanation is that the greater difficulties involved in integrating phase II of the experiment into the schedule have caused a change of attitude among the student teachers. The variation in the number of lessons lost was namely greater in 1970 than in 1969. (The student teachers of 1970 who participated also considered their absence from their training to be more serious than those of 1969 had done.)

Student teachers who participated in the experiment

The answers received from the student teachers who participated in the experiment can be summed up as follows:

1. All student teachers found participation in the experiment enjoyable.
2. All student teachers considered it to be valuable experience to participate in such experiments during their teacher training.
3. All student teachers considered that continued research into the use of CCTV/VR techniques in teacher training should be carried out on a large scale at the School of Education.

To summarize, it can be said that the internal validity of this study is guaranteed by the factorial design presented above. That the experiment should have internal validity is a minimum requirement, without which one cannot even start to interpret an experimental study. It has also been pointed out that the possibility of generalizing the results of the experiment to the population in question is partly dependent on the ecological validity of the experiment.

In conclusion, the results of the attitude questionnaires show that the ecological validity of the experiment is reasonably good. The experiment has not been felt to interfere unduly with the teacher training program. Moreover, the general attitudes towards the experiment and self-confrontation via CCTV/VR have been positive.

32.3 Various sub-studies

In addition to the factorial design for examining the self-assessment of the student teachers, the arrangement of the study also permits investigation of the following problems:

1. *Assessment of student teachers by educational experts*

The micro-lessons of the student teachers have been assessed by four educational experts. These independent assessments have been examined for agreement between two of them. An analysis of variance model was then used to help find out whether, taking the "average assessment" of the experts as a criterion, participation in the experiment had resulted in any demonstrable effects upon the teaching behavior of the student teachers.

2. *The self-assessment of the student teachers and the assessments of the educational experts*

An important goal in teacher training is to develop the skill of the student teachers in interpreting educational processes "objectively", i.e. realistically. In order to study the "degree of objectivity" in the perception and evaluation of the student teachers, it is necessary to have an external criterion. In the experiment this external criterion consists of the "average assessment" of the educational experts. An examination of the "objectivity" of the student teachers' self-assessment, as defined below, also requires, however, that there should be a basis for assessment (e.g. video-recorded teaching situations) and a rating scale (categories) that are identical for both student teachers and experts. The operational definition of "objective" perceptions and evaluations used in this study is based on the experts' average assessment as criterion of the objectivity. As a measure of the deviation of the student teachers' self-assessment from this criterion, the differences have been calculated. This means that a large difference value indicates low objectivity and a small difference value indicates high objectivity in the student teachers' self-assessment. (The expression "*deviations in objectivity*" refers to variations in the calculated difference values.) The way in which the difference between the experts' average assessment and the student teachers' self-assessment varies as a result of the experimental treatment has been studied by means of an ANOVA (Analysis of Variance).

Furthermore, the observations of the educational experts and the student teachers have been studied with a view to finding possible similarities in struc-

ture. The possible occurrence of an overlapping structure between the self-assessment of the student teachers and the average assessment of the experts has been investigated by canonical correlation analyses.

3. *The influence of student teachers' predispositions and personalities on their own perception and evaluation of teaching process mediated by CCTV/VR techniques*

In order to be able to study a possible connection between the student teachers' special perception and evaluation tendencies on the one hand, and aspects of their personality on the other, a battery of group test has been administered, containing different personality tests, cognitive tests and attitude tests. For a detailed presentation of the test battery, see Chapter 8.2.6.

4. *Follow-up studies*

The student teachers were asked to assess the video-recorded micro-lessons again, first six weeks and then four terms after the experiment had been concluded. The purpose of this follow-up was to examine to what extent the teacher training had had any effect on their perception and evaluation of the micro-lessons video-recorded during their second term at the School of Education. An analysis of the special studies in points 3 and 4 has not been included, however, in the present phase of reporting.

33. Assessment and evaluation schedule F III: Validity and reliability

The great majority of studies using CCTV/VR techniques as an instrument for research and training have relied on well-known test methods. If one wants to find the answer to a particular problem, however, it is not usually possible to apply old tests to the new problem.

The development of the measuring instrument "assessment and evaluation schedule F III" (cf. Bierschenk, 1972 f and part 1) started with a preliminary experiment (spring term 1968). The construction of the measuring instrument was based on the following question: What do the student teachers really tell us when they are confronted with their own teaching performances by means of CCTV/VR? Thus the measuring instrument was developed from scratch.

The comments made by the student teachers in the experiment during the process of self-confrontation were recorded on tape and then subjected to content-analytical treatment. This treatment could be carried out with a fairly satisfactory coder agreement, both in deciding the coding units ("information units") in the comments of the student teachers and in coding the information units in accordance with a system of categories (cf. Bierschenk, 1972 f). Thus this treatment resulted in an acceptable coder agreement in the categorization of the student teachers' spontaneous and simultaneous comments during the playback of the video-recorded micro-lessons. The separate categories have then been re-formulated into the statements which make up the assessment and evaluation schedule F III. The results presented in this report are based almost exclusively on this instrument. The statements included in the measuring instrument thereby reflect the problem areas on which the student teachers themselves have spontaneously focused attention. The content analysis only makes it possible to judge the content validity of the instrument, but this is a very essential and desirable form of validity. It is often missing from psychological tests on personality, or is only said to exist on flimsy grounds. It is logically impossible to define "the whole process of perception" and since we cannot specify the process of perception completely, we cannot know whether or not we have a measuring instrument of "true validity". We must, therefore, be content to estimate the validity indirectly.

The problem areas which emerged from the student teachers' simultaneous comments during the process of self-confrontation have been categorized according to the following six dimensions constructed a priori:

1. ego-ego relation
2. ego-pupil relation
3. ego-NPO relation (i.e. relation between ego and non-personal objects)
4. pupil-ego relation
5. pupil-pupil relation
6. pupil-NPO relation

These dimensions are defined in the assessment and evaluation schedule F III by a total of 79 statements. The instrument has been constructed to quantify the individual's ability to (a) assimilate information (perceive) (b) digest information (evaluate) and (c) use this information (modify).

If one wishes to prove the existence of symmetrical relations or to be able to predict a particular type of behavior, the measuring instrument must give evidence of being a reliable measure. A certain amount of information about the reliability of the self-assessment of the student teachers can be gained from an examination of the communality of the variables (h_i^2), i.e. the squared multiple correlation, R_{ik}, can be used as an estimation of the lower limit of reliability of a particular item. If the decision as to whether a certain item has unsatisfactory common variance is based on the criterion $h_i^2 < .30$, then four statements within the area of perception (aspect a_1) and three statements within the area of evaluation (aspect a_2) fail to fulfil this limit value.

Since a separate ANOVA has been carried out for each individual subject-object relationship, summed up over the entire variable domain, an attempt was also made to calculate the reliability of the individual variable domain by means of "Cronbach's alpha". The result indicated, however, that in this case this coefficient is not a suitable measure of reliability. Therefore, the following have been presented as comparison criteria:

(1) the squared average multiple correlation
(2) the average communality and
(3) an estimation of homogeneity according to Spearman—Brown's "prophecy" formula.

Reliability can be defined and estimated in many different ways and for that reason estimation based on a single method can easily lead to over-hasty conclusions.

The examination of reliability has established that, as a whole, the reliability of the student teachers' self-assessment is satisfactory. The reliability of the

individual statements in the educational experts' assessments has been estimated by means of the intra-class correlation coefficient, r_{21}, i.e. for two experts and one teaching occasion. The reliability of the summation variables was also calculated with "Cronbach's alpha".

A closer examination of the reasons for the unreliability of individual items showed that the decision as to the reliability of a particular item could not be based solely on r_{21}. Both m and s and the raw data indicated that for some items there was no response variation at all. In such cases a reliability measure based on variance cannot be used to indicate the agreement in the experts' estimations. For the same reason, "Cronbach's alpha" does not appear to be altogether suitable as a measure of reliability, even if satisfactory reliability coefficients could be demonstrated for certain variable areas (because of larger distribution values).

In summary, it may be said that the low reliability values obtained from the different reliability measurements based on variance have often been caused partly by a lack of variance within certain variable areas, and partly by the low item covariance. The relatively extensive examination of reliability has shown that the assessment and evaluation schedule F III can be regarded as being essentially a reliable measuring instrument.

34. Analyses of results

Data for the complete design exist for the assessment and evaluation schedule F III. This schedule has been made the primary subject for analysis and these observation data have been analysed for both level and structure. The individual analysis programs have been described in Chapter 10. The summary and discussion of results presented in this report concern above all:

1. ANOVA treatment of the student teachers' self-assessment (part 2),
2. ANOVA treatment of the assessment by educational experts (part 3), and
3. ANOVA treatment plus a canonical correlation analysis of the student teachers' self-assessment *and* the educational experts' assessments (part 4).

Problems of research method were discussed in comparative detail in connection with the individual result analyses. Completed result analyses that have not yet been reported include a number of factor analyses and simultaneous comments which have been coded.

If any form of inference statistics is used in the analysis of behavioral observation data, then (1) the prerequisites demanded by a particular statistical mathematic model should be fulfilled, and (2) the precision and power of the statistical tests used should have been explicitly determined.

As omnibus tests, significant F tests are very useful indicators of systematic differences among cell means, but only a careful examination of detail will make it possible to interpret the experimental results thoroughly. If, in addition, the design is rather complex, a large number of F tests are needed, and that in its turn increases the probability of a certain number of tests resulting in random significances. For this reason, one should avoid attaching too much importance to isolated results. The guideline followed in the evaluation of the analyses of results has therefore been the interpretability of the patterns in the F tests. In order to obtain additional and more objective indicators as to whether there is any point in a more thorough interpretation of the main and interaction effects respectively or in carrying out contrast analyses and in commenting on simple effects, the precision and power of the significant F tests have also been calculated.

Thus the individual ANOVA results have been evaluated step-wise. First (1), the interpretability of the patterns in the F tests was examined. Then the precision and power were estimated in order to decide (2.1) the size of the effects and (2.2) the probability of discovering an effect of a particular size. Contrast analyses (3) were not carried out until this point. Only effects at the level of $\alpha = .001$ which show a probability of at least .70 of discovering a particular size of effect have been interpreted, however. Values lower than this are of little use as evidence or as a basis for interpretation.

34.1. Analysis of levels

34.1.1 Step 1: Patterns in the F tests

The interpretation of the first step in the individual analyses of results (parts 2, 3 and 4) shows that the F tests in all three parts of the analysis have led to interpretable patterns. A summary of the patterns in the F tests, referring to the respective part of the analysis, is given in Table 76. Owing to the construction of the F tests as omnibus tests, the interpretation cannot go further at this stage of the analysis than to establish that there is a systematic pattern in the F tests, that in addition permits an interpretation that is meaningful from the point of view of educational psychology.

The pattern in the F tests for the self-assessment of the student teachers shows no significant main effect either in Factor T (self-confrontation via CCTV/VR) or in Factor H (dyadic confrontation in the form of traditional tutoring). In addition, H_0 is accepted for the factor combination TH.

The demonstrable interaction effects partly imply, however, that externally mediated self-confrontation via CCTV/VR leads to reactions that are influenced by the predispositions and predictions of the student teachers and that these reactions are modified by the information that the student teachers have received. It is also worth noting that the student teachers, regardless of the type of influence, seem to modify their perceptions and evaluations from lesson to lesson as far as their own behavior towards an object is concerned (ego dimension). There is no such modification, however, in the subject-object relationships where the student teachers must relate the actions of other persons to their own person (pupil dimension).

With regard to the summation variables, the pattern in the F tests for the average assessments of the experts shows with respect to the TH and ATH interaction effects a greater degree of homogeneity and more significances than the pattern in the F tests for the self-assessment of the subjects. Moreover, the average assessment of the experts in the pupil dimension (pupil-ego, pupil-pupil relationships) has led to significant main effects in Factor H.

An examination of the influence of traditional tutoring with regard to the interaction effects (Factors U, H) shows that these interaction effects (UH, AH) largely apply to the pupil dimension. One possible explanation is that as a result of the tutor's influence the climate of behavior both between student teacher and pupils and among the pupils themselves have changed. But these effects can also simply be a consequence of the presence of the tutor in the classroom, which can have had a subduing effect on the pupils. In this context, it is not possible to decide which of these explanations is the most probable.

As far as the interaction effects involving Factor U are concerned, the analysis of the self-assessment of the student teachers shows that these effects

Table 76. Summary of the significant F tests for the summation variables in the respective sub-analyses

Source	1 Ego-ego relation			2 Ego-pupil relation			3 Ego-NPO relation		
	1	2	3	1	2	3	1	2	3
T									
H									
TH		**	*	*	**				
U	**		**	**	*		**	**	
UT									
UH									
UTH									
A	**	**		**	**	**	**	**	
AT			*				*		**
AH									
ATH		**	*	*	**	**		**	
AU	**						*		*
AUT							**		**
AUH								*	
AUTH									

Source	4 Pupil-ego relation			5 Pupil-pupil relation			6 Pupil-NPO relation		
	1	2	3	1	2	3	1	2	3
T									
H		**			**				
TH	*	*	*					**	
U		**	**		**				
UT		*							
UH		**			*	**			*
UTH									
A	**	**	**	**	**		**	**	
AT									
AH		**		**	**		*		
ATH	*	**	**	*		*	**	**	**
AU	**		*						
AUT				**					
AUH							**		**
AUTH									

1 : Analyses of student teachers' self-assessment
2 : Analyses of educational experts' assessments
3 : Analyses of student teachers' self-assessment and educational experts' assessments
** $F_{.99}(1, 92) = 7.08$
* $F_{.95}(1, 92) = 4.00$

are also mainly restricted to the pupil dimension. Within the ego-NPO relationship both student teachers and experts appear to have observed changes between microlessons 1 and 2.

The significant interaction effects for the factor combinations TH and ATH imply that the combination of T and H produces demonstrable differences between the groups. No effect can be seen in the UTH interaction, however.

The pattern in the F tests has been examined for variations in the objectivity of the student teachers' self-assessment. The interaction effects involving CCTV/VR imply that the processes of self-confrontation have resulted in demonstrable variations in the differences between the student teachers' self-assessment and the experts' average assessment. These effects indicate that there is a variation in the objectivity of the student teachers' assessment of the relations in which ego is the subject. Demonstrable variations in the differences between the assessments of the student teachers given traditional tutoring in the form of dyadic confrontation and those not given this treatment, however, relate to the pupil dimension.

Moreover, the effects in Factor U show that, regardless of the factors T and H, the objectivity of the observations of the student teachers have been influenced by the micro-lessons.

Thus, on the question of the ego-ego relationship (the student teacher's own person) and the behavior of the pupils towards the student teacher (e.g. the extent to which the pupils are "socially provocative"), the difference between the assessment of the student teachers and the experts average assessment shows noticeable variations from micro-lesson 1 to micro-lesson 2. It should be noted that both effects relate to the variable domains where ego is the object.

For Factor A the F tests show a very homogeneous pattern for the observation data of both student teachers and experts (cf. Table 76, sub-analysis one and two). This means that, regardless of the experimental conditions and specific statements, the perception (a_1) and the evaluation (a_2) are different. The homogeneous pattern that appears in all the variable domains is an expected outcome. This factor was included in the analysis as a precision factor and not because of any wish directly to compare perception and evaluation.

In part three of the analysis, Factor A shows demonstrable effects within the ego-pupil relation and within the pupil-ego relation. This result implies that there are differences in objectivity between the perception and evaluation of the student teachers (regardless of the experimental conditions or lesson concerned) within these sectors.

It is not possible to make any deeper interpretation in this step of the analysis. Only an examination of the significant F tests by means of contrast analyses can provide information as to what has caused these significances.

34.1.2 Step 2 of analysis: Precision and power in the F tests

Summaries and discussions of results have been reported in more detail in parts 2, 3, and 4. In these discussions, however, the effects have only been taken into consideration when precision and power estimations have indicated that there is a conclusive basis for interpretation. The predictor variance ($\hat{\omega}^2$) in the significant effects proved to be relatively low in all three parts of the analysis. But since the numerical size of $\hat{\omega}^2$ is dependent on how many sources of variation are included in an ANOVA, a comparison criterion is needed if one is to be able to decide on an objective basis whether the $\hat{\omega}^2$ values presented here are indications of unimportant correlations or perhaps of important empirical results. J. Cohen's (1969) "effect size index" (f), which gives a certain effect size (ES) when all other effects in the analysis have been held constant, has been used for this purpose. Cohen, J. (1969, p. 278) denotes a small effect with $f = .10$. A medium effect corresponds to $f = .25$ and a large effect to $f = .40$.

If f is calculated, Cohen's tables (pp. 282—347) can also be used to decide the power of the F tests. The tables must be used with a certain amount of care, however, since in factorial designs df in the denominator no longer agrees with the n values stated in the appropriate table. In factorial designs there is usually a lower number of df in the denominator than the number stated for the table values. As a result, above all as far as the main effects are concerned, the power that can be read in the tables is an over-estimation of the probability of the effect in question. Main effects with an $f < .35$ constitute an uncertain basis for interpretation (for a more detailed discussion, se Ch. 15.2).

A summary of the probability values (g), referring to the respective part of the analysis concerned, is given in Table 77.

Information on the precision values of the effects ($\hat{\omega}^2$, f) can be obtained from parts 2, 3, and 4. With these estimations of power it becomes possible not only to state that an effect of a particular size does exist, but also to state the degree of probability that this effect really is of the stated size. Table 77 shows how the step-wise procedure has finally produced only a few effects within each separate sub-analysis that are suitable for more detailed consideration and interpretation. In the light of this result, there has been no reason to undertake a more thorough interpretation of the contrast analyses.

34.1.3 Step 3: Post-hoc comparisons

Since it can be difficult to define what is to be regarded as really valuable information, the controls described above were applied in order to decide whether or not a more detailed examination and discussion of the results of the experiment would be worthwhile.

As was pointed out (Chs. 3 and 11), there have unfortunately been far too many cases in which no attempt has been made to state to what extent

Table 77. Summary of power values for the significant F tests ($\alpha = .01$) in the three parts of the analysis (summation variable)

Source	1 Ego-ego relation 1	2	3	2 Ego-pupil relation 1	2	3
T						
H						
TH		.62			.62	
U	> .99		.78	.78		
UT						
UH						
UTH						
A	> .99	> .99		.72	> .99	> .99
AT						
AH						
ATH		.66			.91	.93
AU	.62					
AUT						
AUH						
AUTH						

Source	3 Ego-NPO relation 1	2	3	4 Pupil-ego relation 1	2	3
T						
H					.72	
TH						
U	.83	.67			.83	.61
UT						
UH					.77	
UTH						
A	> .99	> .99		> .99	> .99	.67
AT			.52			
AH					.67	
ATH		> .99			> .99	.55
AU				.57		
AUT	.81		.88			
AUH						
AUTH						

the requirements for a given method of analysis have been complied with, or to decide the precision and power of the tests in question. Only after such estimations have been made, however, is it possible to judge the usefulness of the significant F values for a more detailed study of the relationship between the simple effects. An evaluation of the experiment's data with the F statistics involves testing the null hypothesis. Rejection of the null hypothesis implies that the set of data in question contains systematic effects. An F test does not indicate the direction of the effects, however, nor does it state the precision of the measurement or the probability of an effect being of a certain size. Seen in the light of the power estimations to be found in Table 77, the self-assessment of the student teachers could not on the whole be used for a detailed analysis of the contrasts, and for this reason the contrast data were presented as an appendix (Bierschenk, 1972 e). Compared to the self-assessment of the student teachers, the average assessment of the experts at least with regard to the TH and ATH interaction effects has resulted in a more uniform pattern in the F tests, in greater precision, and in a higher degree of probability for the proven effects. This type of result was expected, however. A large proportion of the variation in the student teachers' self-assessment can probably be traced to differences between the individuals that existed prior to the experiment. Since the analysis of the student teachers' observation data is based on $n = 96$, while the analysis of the average assessment is based on $k = 2$, the standard deviation of the means is smaller in the experts' observations. This in its turn means that the differences between the cell means need not be as large as for the student teachers in order to produce demonstrable effects.

Since there are always deficiencies which could have led to small effects and low probabilities, an account of the contrast analyses may be of interest for further research work. Any reader who is also interested in result analyses 2 and 3 can refer to Bierschenk, 1972 e. Relatively few significances have shown a satisfactory power in the effects and these are therefore discussed in the chapter "Final discussion" (see parts 2, 3, and 4).

In order to obtain a more surveyable perspective of the tendencies that seem to appear in the separate analyses, a summary of the main effects is given in Table 78.

Table 78 shows how the average assessment of the experts alone resulted in demonstrable main effects in Factor H of the experiment. The mean values indicate a more positive assessment of the group with dyadic confrontation.

The self-assessment of the student teachers has for Factor U resulted in three demonstrable effects, all of which involve the *ego* dimension. According to the average assessment of the experts, however, the demonstrable effects primarily involve the pupil dimension. But both student teachers and experts have reported changed values in variable domain 3, i.e. in the re-

Table 77. (Cont.)

Source	5 Pupil-pupil relation			6 Pupil-NPO relation		
	1	2	3	1	2	3
T						
H		.78				
TH					.52	
U		.46				
UT						
UH			.84			
UTH						
A	> .99	> .99		.95	.54	
AT						
AH	.52	.89				
ATH				.76	.73	.93
AU						
AUT	.55					
AUH				.69		.85
AUTH						

1 : Analyses of student teachers' self-assessment
2 : Analyses of educational experts' assessments
3 : Analyses of student teachers' self-assessment and educational experts' assessments

lation between the student teachers and the aspects concerning teaching method. In both cases, the change was positive. In this factor, significant variations in the objectivity of the student teachers' self-assessment can be demonstrated in two cases. In the ego-ego relationship the difference is greatest in connection with lesson 1 and diminishes strongly in connection with lesson 2. The same tendency can be observed in variable domain 4.

Factor A represents two aspects of the measuring instrument, namely perception (a_1) and evaluation (a_2). Differences between perception and evaluation appear in each variable domain for both the self-assessment of the student teachers and the average assessment of the experts. The tendency in both sets of data is the same. The mean values seem to indicate a tendency for the perception of the experts to be more positive than that of the student teachers.

Variations in the differences between the self-assessment of the student teachers and the objectivity criterion are significant only for variable domains 2 and 4. There is a very slight difference between the objectivity criterion and the student teachers' perception regarding the pupil-ego relation (4). No socially provocative behavior on the part of the pupils seems to have occurred. In their evaluation, however, the deviation is comparatively great. The student

teachers evaluate possible behavior of this kind as being rather distressing, while the average assessment of the experts is that it is relatively easy to deal with.

Table 78. A summary of the mean values of the significant main effects for the partial analyses 1, 2, and 3

Variable domain	Student teachers' self-assessment (1)		Experts' average assessment (2)		Student teachers' self-assessment in relation to experts' average assessment (3)	
	Factor H		Factor H		Factor H	
	h_1	h_2	h_1	h_2	h_1	h_2
4	–	–	(5.21	5.14)	–	–
5	–	–	(4.85	4.61)	–	–
	Factor U		Factor U		Factor U	
	u_1	u_2	u_1	u_2	u_1	u_2
1	4.60	4.80	–	–	(.43	.28)
2	(4.64	4.74)	–	–	–	–
3	(4.90	5.04)	(4.95	5.01)	–	–
4	(4.93	5.00)	(5.21	5.15)	(.27	.14)
5	–	–	(4.80	4.66)	–	–
6	–	–			–	–
	Factor A		Factor A		Factor A	
	a_1	a_2	a_1	a_2	a_1	a_2
1	4.90	4.50	5.22	4.90	–	–
2	(4.77	4.62)	5.04	4.59	.28	.03
3	4.66	5.28	4.68	5.28	–	–
4	6.12	3.82	6.15	4.20	(.03	.38)
5	4.88	4.17	5.11	4.33	–	–
6	5.27	5.62	(4.85	5.03)	–	–

() Uncertain basis for interpretation.

34.1.4 Implications of ANOVA results

The use of different analysis techniques and the stepwise approach used in reporting the results aim at making the evaluation more critical and thereby more objective.

We hope that the result analyses can serve as examples of the application of principles of research method and of how one can explicitly prove if and to what extent the assumptions of the statistical-mathematical models are fulfilled. In this way the inferences become meaningful.

In addition, it is hoped that the detailed description of the experimental conditions (Chs. 4 and 5) will provide:

1. increased knowledge of possible ways of using CCTV/VR as a research and training instrument,
2. increased possibilities of repeating behavioral experiments or at least
3. increased opportunities for comparing individual research results.

Finally, keeping in mind Stickell's (1963) examination of research reports concerning "televised and face-to-face instruction", the results can be seen as

4. a contribution to improving the quality of research results dealing with the use of CCTV/VR in educational contexts.

From an empirical point of view this experiment has produced results which cannot be made the basis of dichotomous decisions, i.e. either-or decisions. If the experimental results are interpreted purely pragmatically or from the point of view of economy, it might seem reasonable simply to recommend the cheapest alternative, i.e. the student teachers seem to need no tutorship in the form of dyadic confrontation and/or externally mediated self-confrontation via CCTV/VR, since the experimental conditions have not led to any demonstrable main effects.

Speaking against such a decision, however, are certain alternative explanations of the null hypothesis and the fact that one cannot draw such far-reaching conclusions from a single experimental result. Another important argument against such a decision is the consideration of the possible long-term effects and this aspect will be examined more closely in the experiment's follow-up studies.

There can be many reasons for accepting the null hypothesis for a particular factor or factor combination, and therefore it is difficult to say anything definite when the null hypothesis is accepted. Despite this, a section entitled "Final discussion" has been included for each part of the analysis, presenting some more or less speculative considerations that could be regarded as possible implications of the results. In the following, we try to summarize point for point. The first figure states the order of sequence. The figure following the colon denotes the part of this study in which the result has been discussed in detail. Thus in the case of the self-assessment of the student teachers, for example, (1:2) states the results of both externally mediated self-confrontation via CCTV/VR (Factor T) and dyadic confrontation in the form of traditional tutoring (Factor H), (2:2) states the results of the teaching situations (Factor U), (3:2) states the results of the aspects in the assessment and evaluation schedule (Factor A), (4:2) states the results of the conclusive higher order interaction effects. After the results given in each sub-section, possible implications are presented.

The educational and psychological implications that have already been presented in connection with the individual experiment results have been divided between the following two categories: (1) Implications based on the

results of the experiment and (2) implications based on more speculative considerations.

The first category covers implications that either (1) are of a descriptive nature, i.e. establishing facts, or (2) are based on experimental data which provide a conclusive basis for interpretation. The second category covers implications that either (1) are based on experimental data which do not provide a conclusive basis for interpretation or (2) are of a purely speculative nature.

The borderline between the two categories, however, can doubtless be challenged in many cases. The main reason for making this division was to make clear which implications can be said to be based on the *conclusive* effects of the experiment. Another motive was that the large number of implications presented could give an undesirable impression of indecision. Finally, some results and implications have been further clarified by short comments.

1:2 *Results:* Student teachers' self-assessment

Neither externally mediated self-confrontation via CCTV/VR nor dyadic confrontation in the form of traditional tutoring (Factor H) have led to significant effects.

Implication, based on these results

The experimental conditions produce no effect, i.e. have not led to any difference in the ability of the student teachers to discriminate.

Implications, based on more speculative considerations

1. The experimental conditions lasted for too short a time for the various influences to achieve observable effects.
2. Self-confrontation requires systematic training in receiving and adapting first-hand information, i.e. "self"-information not mediated verbally.
3. Self-confrontation entails a temporary de-organization or de-automatization, the first phase of which produces in many people feelings of surprise, fear, shock and/or the adoption of defensive attitudes.
4. The tutor has not succeeded in influencing the student teachers to any degree since they have not yet developed suitable test criteria, i.e. educational-psychological norms.
5. Tutor and student teacher avoid a relevant critical analysis by, e.g. using words such as "a verbal portrait of an individual" (Stoller, 1970, p. 11) in order to avoid having to make a critical examination of the student teacher's own behavior.

1:3 *Results:* Educational experts' assessments

The null hypothesis is accepted for externally mediated self-confrontation via CCTV/VR (Factor T), but is rejected for dyadic confrontation in the form of traditional tutorship (Factor H). Significant effects have been demonstrated within pupil-ego relations (4) and pupil-pupil relations (5).

The estimation of precision and power indicates, however, that these effects can hardly be regarded as an acceptable basis for interpretation.

Keeping in mind the far from conclusive effects, a few possible but rather more hypothetical interpretations are presented below.

Implications, based on more speculative considerations

1. The tutor has influenced the student teachers in such a way that their behavior has become more positive. Thus the tutor has successfully mediated both his teaching strategy and some concrete suggestions for action and the student teachers have succeeded in modifying their own behavior on the basis of the dyadic confrontation. In addition, the result seems to be in agreement with the tutor's intention, namely to focus the tutorship upon problems of pupil activation.
2. A prerequisite of traditional tutorship is the presence of the tutor during the actual teaching process, and there is therefore a possibility that these effects have arisen as a result of the subduing effect that the teacher's presence has had on the pupils' activity.

The second alternative (2) seems the most probable, since none of the relations where the student teacher is the subjet have led to demonstrable effects.

1:4 *Results:* Student teachers' self-assessment and educational experts' assessments

Neither externally mediated self-confrontation via CCTV/VR (Factor T) nor dyadic confrontation in the form of traditional tutorship (Factor H) have led to significant effects.

Implication based on these results

Since no deviations in the differences between the student teachers' self-assessment and the average assessment of the experts can be demonstrated, the objectivity, e.g. the "gap" between the objectivity criterion and the student teachers' self-assessment has not been influenced by the experimental treatment. In the context of the definition of objectivity applied in this study, the deviations in objectivity in the observations of the student teachers do not vary as a consequence of the experimental treatment.

2:2 *Results:* Student teachers' self-assessment

The teaching situations, i.e. micro-lessons 1 and 2 (Factor U), have resulted in significant effects concerning the ego dimension: ego-ego relation (1), ego-pupil relation (2), ego-NPO relation (3). For the pupil dimension (variable domains 4—6), the null hypothesis is accepted. The estimations of precision and power indicate that the effect within the ego-ego relation may be looked upon as a conclusive basis for interpretation.

Implications, based on these results

1. Taking micro-lesson 1 as a starting-point, the student teachers seem to be able to predict their own behavior and test these predictions during micro-lesson 2, and subsequently to modify the structure of perception and evaluation, or the concrete behavior.

   As far as the student teachers' perception and evaluation of the pupil dimensions is concerned, the observations indicate no changes.

Implications, based on more speculative considerations

1. The pupils' behavior has not changed (pupil as subject).
2. The student teachers lack criteria for assessing the behavior of the pupils.
3. The student teachers have been primarily occupied with their own person and have therefore not had time to study the behavior of the pupils in any detail.

2:3 *Results:* Educational experts' assessments.

For the teaching occasions (Factor U) the null hypothesis is rejected within the ego-NPO relation (3), pupil-ego relation (4) and pupil-pupil relation (5). The mean values indicate a positive change with regard to the ego-NPO relation, while the change for the other relations (4, 5) is negative. Estimations of precision and power, with the possible exception of the pupil-ego relation (4), indicate an inconclusive basis for interpretation, and for that reason the interpretations suggested below should be regarded as being hypothetical.

Implications, based on more speculative considerations

1. The examination of effects implies that the behavior of the pupils becomes more disturbing during lesson 2, despite the fact that the student teacher's teaching technique improves. This could mean that there is a "relaxation effect". It is possible that the first micro-lesson was rather tense, since neither the pupils nor the student teacher were accustomed to the situation, while the second lesson could have been felt by both student teacher and pupils to be a more "everyday" situation, i.e. the behavior of the pupils has been more normal, in other words "more disturbing".
2. The small change in quality of the student teacher's teaching method has little significance for the teacher-pupil relation or, in other words, for the climate in the classroom.

2:4 *Results:* Self-assessment of student teachers and average assessment of educational experts

For the teaching occasions (Factor U) as cause of variation, the null hypothesis was rejected for the ego-ego relation and the pupil-ego relation. The mean values of the cells indicate variations in the objectivity of the student teachers' self-assessment that were greatest in micro-lesson 1, but diminished substantially in micro-lesson 2. An examination of the precision and power in this result implies that there are no conclusive empirical grounds for interpretation of the effects.

Implications, based on more speculative considerations

1. The diminished deviation from the objectivity criteria of the student teachers' self-assessment from micro-lesson 1 to micro-lesson 2 implies a more realistic assessment of their performances.
2. From the point of view of the student teachers, it seems obvious that progress has been made on the second teaching occasion. According to the educational experts, however, this is not the case.

3:2 *Results:* Student teachers' self-assessment

The perception and evaluation of the student teachers are included as Factor A (aspect) in the analysis of variance. The null hypothesis is rejected for this factor within all six subject-object relations. The assessments of precision and power show very high values, with the exception of the ego-pupil relation. This means that the effects can be regarded as providing a conclusive basis for interpretation.

Implications, based on these results

1. Irrespective of the experimental conditions, the student teachers' perception differs from their evaluation. Moreover, Table 78 shows that their perception has resulted in positive scores.
2. Irrespective of the experimental conditions, the student teachers have evaluated these behavioral aspects as being essential and undisturbing, with the exception of the pupil-ego relation where the student teach-

ers indicate that "socially provocative behavior" (if it had occurred) would have been considered relatively distressing.

3. A comparison of the student teachers evaluation, which relate to their perception of the behavioral aspects constituting variable domains 4 and 5, imply that the student teachers appear to have a high level of tolerance when it comes to the behavior of the pupils towards each other (variable domain 5), even if it is felt to be comparatively undisciplined, while direct action on the part of the pupils against the student teacher (with a conscious or unconscious element of provocation in it) is felt to be distressing.

3:3 *Results:* Educational experts' assessments

The experts' perception and evaluation (Factor A) show significant effects in all six subject-object relations. With the exception of the pupil-NPO relation, the estimations of precision and power have led to very high values. In this context, the effects can safely be regarded as a satisfactory basis for interpretation.

Implications, based on these results

1. Irrespective of the experimental conditions, the experts' perception differs from their evaluation. The experts' perception has in each variable domain resulted in positive scores.
2. Irrespective of the experimental conditions, the experts have in each case evaluated these behavioral aspects of the situation in question as being both essential and undisturbing (cf. 3:2).

3:4 *Results:* Student teachers' self-assessment and educational experts' assessments

With regard to the objectivity of the perception and evaluation of the student teachers (Factor A), the null hypothesis is rejected within ego-pupil (2) and pupil-ego (4) relations. Table 78 shows that the student teachers' perception of the ego-pupil relation differs negatively from that of the experts, while there is only a slight positive deviation in the evaluation. Within the pupil-ego relation, the situation is the exact reverse. The estimations of precision and power show, however, that only the effect within the ego-pupil relation provides a conclusive basis for interpretation.

Implication, based on these results

1. Irrespective of the experimental conditions, the deviation in the objectivity of the student teachers' perception of their own actions towards the pupils (2) is negative.

Implication, based on more speculative considerations

1. The effect size for the pupil-ego relation can hardly be said to indicate a conclusive basis for interpretation. While the student teachers evaluate socially provocative behavior as rather distressing, the experts evaluate it as being comparatively undisturbing. In his role as a leader, the teacher has a decisive influence on the social-psychological structure. His interpretation of what is socially provocative behavior should, therefore, be highlighted to a greater extent in educational contexts.

Externally mediated self-confrontation via CCTV/VR (Factor T) *and* dyadic confrontation in the form of traditional tutorship (Factor H), i.e. the factor combination TH, have only led to demonstrable effects in connection with the experts' average assessment. For the TH interaction the null hypothesis is rejected within the ego-ego relation (1), the ego-pupil relation (2), and the pupil-NPO relation (6). A possible explanation of the effects in the TH interaction is that the experiment was carried out in two phases, with a year's interval, which could have caused a change in the experts' perception and evaluation structure. The estimations of precision and power imply, however, that these effects should not be made the basis of any interpretation.

The experimental factors T and H plus the factor combination TH have also been examined for signs of interaction with Factor A (aspect) and Factor U (teaching occasion). A short summary is given below of results that fulfil our criteria for interpretation.

4:2 *Results:* Student teachers' self-assessment

Only a few of the interaction effects fulfil the requirements. The following interactions have been examined in more detail: (1) AUT within the ego-NPO relation, (2) ATH within the pupil-NPO relation and (3) AUH within the pupil-NPO relation.

Externally mediated self-confrontation via CCTV/VR (Factor T) in combination with aspect (Factor A) and/or teaching occasion (Factor U) led to changes within the ego dimension (variable domains 1—3). On the other hand, dyadic confrontation in the form of traditional tutoring (Factor H) in combination with Factor A and/or U resulted in changes within the pupil dimension (variable domains 4—6).

1. When considering the AUT interaction the perception of student teachers receiving only externally mediated self-confrontation via CCTV/VR shows no change. There is, however, a tendency for the evaluation to become more positive.
2. The perception of the student teachers who were not given this treatment showed a positive change, while the group's evaluation appears to be relatively unchanged.

3. When considering the ATH and AUH interaction the perception of the student teachers who received only dyadic confrontation is demonstrably more positive than the perception of those not given this treatment. However, through dyadic confrontation the student teachers' perception became demonstrably more negative.
4. The perception of the student teachers who were not given this treatment showed a positive change.
5. The evaluation reflects the tendency for the evaluation of the student teachers receiving traditional tutoring to change positively, while the evaluation of those not influenced in this way changes negatively. No significant differences between the simple effects for points 1, 2, 4 and 5 were demonstrable.

Implications, based on these results

1. It seems reasonable to assume that above all the student teachers receiving only externally mediated self-confrontation via CCTV/VR should have shown a change in perception, while the perception of those not receiving this treatment should have been more constant. An unexpected tendency is for the evaluation to change as a result of the self-confrontation via CCTV/VR, despite the fact that the evaluation structures are possibly more difficult to influence. This effect must naturally be looked upon as a relatively isolated result, but nevertheless it seems to point in the same direction as the result obtained by Perlmutter et al. (1967, pp. 900—905), namely that the subjects incline to change their structure of evaluation first.
2. The fact that the perception of the student teachers receiving only traditional tutoring (dyadic confrontation) changes negatively can possibly depend upon the fact that the tutor has pointed out to them certain concrete aspects of their behavior while at the same time verifying their evaluations.

4:3 *Results:* Educational experts' assessments

Several of the interaction effects fulfil our requirements for interpretation. However, Factor T in combination with Factor A and/or U has neither within the ego dimension nor within the pupil dimension resulted in demonstrable interaction effects. Factor H in combination with Factor A and/or U has, on the other hand, only led to significant effects within the pupil dimension. With respect to our criteria for interpretation, the following interaction effects have been examined in more detail: (1) UH within the pupil-ego relation and (2) AH within the pupil-pupil relation.

Furthermore, the ATH interactions in the variable domains 1, 2, and 6 show medium and large effect sizes. But since the educational experts'

assessment of the protocol material was carried out in two phases (first phase 1969, second phase 1970), it cannot be precluded that part of the variance may be traceable to a change in the experts' structure of perception and evaluation. With regard to the UH and AH interactions, the results can be summarized as follows:

1. The educational experts assess the teaching performance of the student teachers who received traditional tutoring on the occasion of lesson 1 to be approximately the same as the teaching performance of those who did not receive this treatment. On the other hand, the student teachers who received no tutoring were judged to be less proficient in their teaching in lesson 2 than those who were given this treatment.
2. On the question of the behavior of the pupils towards one another, their behavior with the student teachers who received traditional tutorship was significantly more disciplined than with those who had not.

Implications, based on these results

1. The tutor has had a positive influence on the behavior of the student teachers.
2. Traditional tutorship requires the presence of the tutor in the classroom, which could have had the effect of subduing the activity of the pupils (cf. 1:3 above).

4:4 *Results:* Student teachers' self-assessment and educational experts' assessments

The majority of the significant interaction effects also fulfil our requirements for an effect size. Both the effect size of (1) the ATH interaction within the ego-pupil relation and (2) the AUT interaction within the ego-NPO relation fulfil these requirements for the ego dimension. For the pupil dimension the requirements are fulfilled by: (1) the UH interaction within the pupil-pupil relation, (2) the ATH interaction within the pupil-NPO relation, and (3) the AUH interaction within the pupil-NPO relation. Hence the objectivity of the student teachers' self-assessment seems to vary significantly in the following respects: Externally mediated self-confrontation via CCTV/VR has led to significant deviations within the ego-NPO relation, while traditional tutoring has led to demonstrable deviations within the pupil dimension (variable domains 4, 5, and 6).

Implications, based on these results

1. The deviation in the objectivity of the student teachers' assessment of the ego-NPO relation increases for both groups from micro-lesson 1

to micro-lesson 2. The simple effects are not significant, however. The results imply that the student teachers expected an improvement in their teaching method in lesson 2, and that this expectation was felt to be born out irrespective of the actual "objective" state of affairs. On the other hand, the student teachers in the group not having externally mediated self-confrontation via CCTV/VR, have not had the same opportunity of seeing their expectations confirmed in the TV monitor.

2. The deviation in the objectivity of the student teachers' assessment of the pupil-pupil relation shows for both groups an increase in the second lesson. If the comments of the tutor are the source of this increase, there seems to be a disagreement in outlook between the tutor and the panel of experts.
3. The deviation in the objectivity of the perception of the pupil-NPO relation by the student teachers receiving traditional tutoring decreases, while that of those not receiving this treatment increases. For the deviations in evaluation the opposite is the case. One possible explanation could be that the TV monitor has had a standardizing effect on perception, while self-confrontation via the CCTV/VR technique has led to greater deviation in evaluation. The simple effects, however, are not significant within the pupil-NPO relation.

To summarize. The account of the results given above shows that the experimental conditions have not on the whole led to main effects that are significant on the chosen level of significance or that can be regarded as constituting a conclusive basis for interpretation. Taking the third analysis of results into account, this means that there are no deviations in the objectivity of the student teachers' self-assessment as a consequence of either traditional tutoring or externally mediated self-confrontation via CCTV/VR.

On the other hand, examination of the interaction between the experimental conditions and the other sources of variation included in the analysis of variance has produced a number of significant effects that fulfil our criteria for interpretation.

The interaction effects that have been examined more closely imply, for example, that traditional tutorship in the form of dyadic confrontation (Factor H) in combination with perception and evaluation (Factor A) and/or teaching occasion (Factor U) has led to changes in the pupil dimension (variable domains 4—6). This result has emerged in both the student teachers' self-assessment and in the average assessment of the experts.

On the other hand, analysis 1 and 3 show that externally mediated self-confrontation via CCTV/VR (Factor T) in combination with perception and evaluation (Factor A) and/or teaching occasion (Factor U) has led to changes within the ego dimension (variable domain 3).

34.2 Analysis of structure

In order that we might study the structural connection between the student teachers' self-assessment and the experts' assessments (average assessment), the data ware treated by means of canonical correlation analysis (cf. e.g. Tatsuoka, 1971). This technique indicates:

1. whether both sets of data are demonstrably related to one another, and
2. the way in which these sets of variables can be combined so that the correlation between the components is at a maximum.

If, as in the present study, it is a question of examining two relatively large sets of variables from the point of view of their interrelationship, then one is primarily interested in a few linear combinations in each group. The variable combinations with the highest correlations are examined first. Moreover, the model means that the structure can as a rule be described almost completely by the first canonical variables, i.e. with a few uncorrelated linear combinations. In other words, the model leads to the relation between the two sets of variables being reduced to its simplest form. For that reason, the method seems to be particularly suitable for use in explorative studies. For a more detailed discussion and description of the method, see part 4, Chs. 28 and 29.

The canonical analysis has been carried out in three stages:

1. the material was examined to find out if there were any significant bivariate relations at all,
2. then the way in which the different variables have contributed to the relation in question was examined, and
3. finally an attempt was made to give these correlations a meaningful content.

Step 1 showed that there are significantly correlated dimensions or common structures in the student teachers' and educational experts' observation data, which are summarized in Table 79.
As can be seen in the table, the analysis has shown nine significant correlated dimensions in the perception structure, while only three correlations have become significant for the evaluation structure. It is obviously easier to achieve a common structure for perception than for evaluation. But for the second lesson, the perception structure shows fewer interrelated dimensions.

Step 2 showed that the weights in the individual dimensions have not on the whole fulfilled the agreement criteria. Thus no common interpretable index can be constructed for either the perception or the evaluation.

Table 79. Number of significant canonical relations (R_c) for perception (a_1) and evaluation (a_2)

Variable domain	Micro-lesson 1 a_1	a_2	Micro-lesson 2 a_1	a_2
1. Ego-ego relation	1	1	1	0
2. Ego-pupil relation	1	0	0	1
3. Ego-NPO relation	2	0	0	0
4. Pupil-ego relation	1	0	1	0
5. Pupil-pupil relation	1	0	1	0
6. Pupil-NPO relation	0	0	0	1

Step 3 showed that the weights within each individual dimension have resulted in different signs, and for that reason no separate and interpretable expert or student teacher indexes could be constructed.

Implications

The results of the canonical analyses carried out on the observation data of the educational experts and student teachers show that there are significantly correlated dimensions in the experts' and student teachers' perception and evaluation. As Cooley and Lohnes (1971, p 169) point out geometrically, the canonical correlation can be interpreted as a measure of the extent to which people occupy the same relative positions in the test space of the first set of variables as they do in the test space of the second set.

Since we are concerned with the relationships between the student teachers' and the pedagogical experts' perception and evaluation respectively, the results in Table 79 show that similarities are evident in the perception area. With respect to the evaluation area, however, there are only three significantly correlated dimensions. This result indicates that there is little association between the student teachers' and the pedagogical experts' evaluation. One consequence of the dissimilarities in structure may be that tutors and student teachers run a considerable risk of misunderstanding one another when they try to discuss separate components of a more complex teaching process.

As far as the research method is concerned, difficulties have also arisen in applying this model to the data in question. Some of the problems are discussed in Bierschenk (1972 c). Another possible approach, which might be more suitable for this particular type of problem, is the development of the canonical correlation analysis model into a hypothesis-testing model, where one only decides upon one weight vector (paired components get the same weights) and then calculates the correlations.

35. Recommendations for further research

The results and implications presented above are based on evaluation of only *one part* of the observation data which have been collected in connection with the experiment described in Chapter 5. In this context the recommendations can be divided into two classes:

(1) general recommendations for further research that can be based on the results and experiences gained during the experiment, and

(2) recommendations for further analysis, i.e. a study of questions that are stated in the project's original formulation of the problem and for which the author has already collected data.

35.1 General recommendations

On the basis of the results of this experiment which have been presented and discussed in a summarized form in this report, the following can be stated:

1. The teaching staff at the Malmö School of Education, the student teachers training as grade 4—6 teachers who did not take part in the experiment, and the subjects of the experiment have all experienced and evaluated the experiment and thereby externally mediated self-confrontation via CCTV/VR as being "of importance" for teacher training. In addition, the subjects consider self-confrontation via the CCTV/VR technique to be "important" for personality development (one of the goals of teacher training). This result is a positive evaluation of both the medium and the technique of self-confrontation.

   It must be considered an extremely important goal for future teacher training to give the student teachers the opportunity of studying and experimenting with their own behavior under systematic and controlled conditions in order to be able to develop different teacher roles and to follow the development of the teaching process.
2. In order to be able to improve the opportunities for research and further investigations into the dimensions which have been studied, the establishment of "*Micro-Lesson Laboratories*" (MLL) is recommended. MLL

would make possible both individualization and an increased number of training lessons.

As has been made clear in the results presented above, the interaction of the experimental conditions (T, H) with the lesson occasions (U) resulted in significant *and* interpretable effects. On the basis of these results, it can be implied that a longer experimental period and more lessons might very well produce a more definite result. If the student teachers were allowed to give a training lesson regularly, every week for example, it would be possible to make a controlled examination of a greater number of variables in the teaching process than the present experiment has permitted.

3. The development of an MLL system would also facilitate a systematic training of student teachers (and other categories of people) in receiving and processing first-hand information, i.e. non-verbally mediated "self"-information. The experimental data imply that micro-lessons, mediated via the CCTV/VR technique, tend to standardize the perception of student teachers and educational experts.
4. MLL should be equipped with a sufficiently large number of video-recorders and video-tapes to make it possible to store video-recorded micro-lessons for a fairly long period of time. This storage would enable student teachers to re-assess regularly the teaching process in question. The experimental data imply that the student teachers who are given the opportunity of seeing their own lessons via CCTV/VR change their evaluation structure. Being able to get the process of evaluation under systematic control must surely be a very essential goal for teacher training in the future.
5. The results of the structure analyses imply that it is important to investigate in more detail the process of evaluation. A study of the development of the student teachers' perception and evaluation structures should be placed in the centre of future research into CCTV/VR techniques, primarily because the null hypothesis for the main effects in the experimental factors has been accepted.
6. A further study should be made of the effects of the tutoring process on student teachers, since it is not possible to establish within the framework of the experiment whether the observed effects depend simply upon the presence of the tutor in the classroom or upon the influence of the dyadic confrontation.
7. Some experimental results appear to confirm the hypothesis that the expectations of the student teachers direct what is observed in a teaching situation and that these expectations via the TV screen are felt to be corroborated whatever the actual "objective" state of affairs is. Keeping this in mind, a closer examination should be made of the degree to which the student teachers' predispositions influence the processes of both perception and evaluation. (The author has data available for study of this question.)
8. In order to increase the validity of the measuring instrument, follow-up studies are needed and an extension of the experiment to schools outside the School of Education.
9. The tutor in the experiment has been *one* lecturer in Methodology. In addition he has a particular interest in educational psychology and cannot, therefore, be regarded as being quite representative of the body of lecturers as a whole. Thus no generalizations should be made from the experimental data as far as lecturers in teaching method are concerned. The way in which MLL techniques could be integrated into the School of Education's courses in teaching method requires further studies.
10. Systematic and controlled studies of personality-psychological and social-psychological dimensions in the teaching process require not only new approaches in research method such as the micro-teaching techniques and the CCTV/VR system, but also new approaches in psychometrics, such as the development of statistical-mathematical models that can deal with complex problems. As is implied in the discussion of the separate parts of the analysis of results, the best experimental design and new methods of observation are of little use if the statistical analyses are unsuitable. Thus work is needed here on the development of new evaluation methods.

35.2 Continued tasks of analysis

The experiment presented above is exploratory, which means that a sizable amount of data has been collected in order that the problem might be studied from several different aspects.

The analysis of the data was primarily based on the assessment and evaluation schedule F III. But the test evaluation and coding work have also been completed for all the other data. These data have been stored on magnetic tape, ready for continued analysis. A few examples of such tasks of analysis are given below:

1. *Analysis of student teachers' comments*

The comments made by the student teachers during the process of self-confrontation were recorded on tape and then worked over in an analysis of content. The coder agreement in coding the physical units, e.g. information units, has been checked.

After this check, recordings were made of both simultaneous comments and comments made during dyadic confrontation, i.e. what the student teachers

and the tutor said during the traditional tutoring. These data have been treated by means of frequency statistics. One of the aims of the evaluation is to examine to what extent the cognition is (1) ego-centered, (2) pupil-centered and (3) topic-centered.

In addition, the analysis of the tutoring comments permits a more systematic examination of the dyadic process of confrontation. The tutor plays a central role in teacher training, and for that reason a more detailed analysis of the pattern of "face-to-face" communication is an important research task.

2. *Analysis of the student teachers' reactions to repeated confrontations with one and the same micro-lesson*

Assuming that repeated confrontation experiences with a single video-recorded lesson should influence the student teachers' perception and evaluation of the teaching process, their micro-lessons have been played back to them three times. The analysis carried out on the assessment and evaluation schedule F III has resulted in an ANOVA, the examination of which is not yet fully completed, however.

3. *Long-term follow-up of student teachers' self-assessment*

At the end of their second term at the School of Education (six weeks after the completion of the experiment) and at the end of their period of training at the School (sixth term), the student teachers have been asked to assess once again the micro-lessons video-recorded during the experiment. The test evaluation and coding work have been completed. In addition, these data are stored on magnetic tape. An analysis of these data would aim at studying the extent to which the teacher training has had any effect on the student teachers' perception and evaluation of the micro-lessons video-recorded during the second term.

4. *Analysis of measuring instrument*

The assessment and evaluation schedule F III has been the main instrument in the experiment but this, like some other schedules included in the group-test battery, is a new construction developed especially for this experiment. To enable the construct validity of the instrument to be studied, a series of factor analyses (cf. Ch. 10) have been carried out for both schedule F III and schedule F II (Identification experiences). As far as schedule F III is concerned, the analytical work has been completed, but it has not yet been reported. An evaluation of schedule F II, e.g. through correlational studies, is also an important task if one wishes to study perception and evaluation tendencies specific for one individual.

5. *Analysis of the influence of student teachers' predispositions on their perception and evaluation of teaching processes mediated via video-recording*

One interpretation of the results presented above has been that the individual's degree of satisfaction with his own performance before seeing the recording decides to some extent what he will observe on the TV screen, in what way he will evaluate it and what changes it will cause in his attitude.

The analyses of results presented above have been carried out with a view to discovering possible differences between the experimental groups. They imply that one also ought to carry out analyses on the level of the individual, e.g. an analysis of the connection between the experimental results and different personality variables. A group-test battery was administered (see Ch. 8.2.6) for the purpose of showing to what extent the student teachers' perception and evaluation of their own teaching was directed or influenced by the individual's (1) cognitive ability, (2) ability to maintain emotional balance, (3) access to adequate social behavior, (4) ability to use pupil-adapted (concrete) language, (5) ability to stimulate and control the teaching process, (6) ability to maintain opinions despite different types of provocation, (7) ability to achieve an integrative behavior, (8) ability to accept himself and others, (9) ability to make perceptual analysis, and (10) ability to maintain a high level of energy and attention.

The evaluation and coding of separate tests and schedules included in this battery have been completed and these data are stored on magnetic tape.

Förkortningar och statistiska symboler

Förkortningar:

16 PF	Personlighetstest: Cattells 16 Personality Factor Questionnaire
ALGOL	Algorithmic Orientated Language. Ett programmeringsspråk som genom sin internationella status och sin stränga logiska byggnad är ett mycket generellt språk
ANOVA	"Analysis of Variance", variansanalys
BMD	"Biomedical Computer Programs"
CD 3600	Control Data 3600 Computer System
CD 8090	Terminalutrustning till CD 3600
$D_{11, 12}$	Dyadisk konfrontation i samband med traditionell handledning, experimentgrupp 1 (första indexsiffran), tillfälle (1, 2)
EK 9	Nio personlighetsvariabler ur EPPS
EPPS	Personlighetstest: Edwards Personal Preference Schedule
ETV	Educational television. Undervisningstelevision som avser program av allmänt informativ natur
exp	Experiment
FA	Formulär: Skriftlig redovisning av upplevelserna omedelbart efter lektionsslut (5 min)
FB	Formulär: Öppna frågor: 6 frågor till subjekt-objekt relationerna (1–6) samt 1 fråga avseende handledning
fb	Fackskolebetyg, slutbetyg från en fackskola
FC	Formulär: Öppna frågor: Sammanfattande värdering av experimentet
F&D	Försöks- och demonstrationsskola
F I	Skattnings- och värderingsformulärets kortare version
F II	Formulär för skattning av identifikationsupplevelser
F III	Skattnings- och värderingsformulär (huvudinstrument)
F IV	Formulär för skattning av elevattityder
F V	Formulär: Forskningsvärdering (deltagare i experimentet)
F VI	Formulär: Forskningsvärdering (ej deltagare i experimentet)
F VII	Formulär: Forskningsvärdering (personal)
F VIII	Formulär: Personlighetstest avsett att mäta självaccepterande och accepterande av andra
F IX a, b, c, d	Formulär: Personlighetstest avsett att mäta självvärdering
FORTRAN	Formula Translation. Ett mera telegramstilsbetonat programmeringsspråk, vars utveckling började inom IBM-företaget och som används för programmering av beräkningsuppgifter
fp	Försöksperson, -s, -en, -ens
fpp	Försökspersoner, -s, -na, -nas
gr	Grupp
hl	Handledare
ht	Hösttermin
IBM	International Business Machines
IPAT	Institute for Personality and Ability Testing
IPO	Icke personella objekt
IPR	"Interpersonal Process Recall"-teknik är en metod för att stimulera återvinning av minnesbilder och upplevelser efter en dyadisk konfrontation. Återgivning av denna konfrontation sker via ljudband- och/eller videobandspelare
IT	Inträdestest. Testbatteri som administrerades vid början av fpp 1:a lärarhögskoletermin
ITV	Intern television. Television förmedlad via kabel
ITV/VR	Intern television och videobandningsteknik
JND	"Just noticeable differences" – ett mått på åtskillnadsminimum
K	Kvinnliga lärarkandidater
LDC	Lunds Datacentral
LH	Lärarhögskola
LHM	Lärarhögskolan i Malmö
lk	Lärarkandidat, -s, -en, -ens
lkk	Lärarkandidater, -s, -na, -nas
LP	Lektionsplanering före respektive mikrolektion (1, 2)
M	Manliga lärarkandidater
M 1	Mellanstadielärarlinje, terminskurs 1
M 2	Mellanstadielärarlinje, terminskurs 2
min	Minut, -er
MLL	Mikro-Lektions-Laboratorier
ml	Mikrolektion, -er som är en nedbantad lektion och som karakteriseras av en mindre komplexitet än en vanlig lektion
$ml_{1, 2}$	Mikrolektion 1 respektive 2 i experimentet
$ml_{11, 12}$	Experimentgrupp 1, mikrolektion (1, 2)
$ml_{21, 22}$	Experimentgrupp 2, mikrolektion (1, 2)
$ml_{31, 32}$	Experimentgrupp 3, mikrolektion (1, 2)
$ml_{41, 42}$	Experimentgrupp 4, mikrolektion (1, 2)
ml_x	Mikrolektion x är en mikrolektion som har utförts av fp II under försöket vt 1968
NPO	Non-personal object
nr	Nummer
O	Allmän beteckning för operation, vilken som helst
Pers test	Personlighetstest. Testbatteri som administrerades vid slutet av fpp andra lärarhögskoletermin
Pr	Praktikvecka under fpp andra lärarhögskoletermin
PÅ	Personlighetstest avsett att mäta åsiktsförändringar
R 10 D	Personlighetstest avsett att mäta dogmatiska tendenser
R 10 F	Personlighetstest avsett att mäta auktoritära tendenser
S	Simultankommentarer under externt förmedlade självkonfrontationsprocesser via ITV/VR
sb	Studentbetyg, slutbetyg från ett gymnasium
SÖ	Skolöverstyrelsen
term	Termin
tim	Timmar
TV	Television
UkW	Ultrakurzwelle
USTM	Intelligenstest, delprov A: uppmärksamhet
VR	"Video tape recording". Videobandning av program (lektioner) på magnetiskt band för autentisk återgivning av undervisningsprocesser
$VR_{11}, \ldots, VR_{110}$	Experimentgrupp 1, videobanduppspelningar (1–10)

$VR_{21}, \ldots, VR_{24}$	Experimentgrupp 2, videobanduppspelningar (1–4)
$VR_{31}, \ldots, VR_{310}$	Experimentgrupp 3, videobanduppspelningar (1–10)
$VR_{41}, \ldots, VR_{44}$	Experimentgrupp 4, videobanduppspelningar (1–4)
VR_{1x-4x}	Experimentgrupp 1–4, videobanduppspelning av ml_x, fpp II (försök vt 1968)
vs	Vändning av bipolära skalor så att den positiva polen av skalan blir 7
vt	vårtermin
WIT III	Av Westrin (1964) utvecklat intelligenstest, innehållande fem deltest: (1) Klassifikationer, (2) Analogier, (3) Motsatser, (4) Sifferkombinationer och (5) Pussel.
X	Symbol som anger en agerande organism, vilken som helst
XOX	Symboler som anger en subjekt-objekt-relation. Den agerande organismen (X) utför en operation mot sig själv (X)
XOY	Symboler som anger en subjekt-objekt-relation. Den agerande organismen utför en operation mot ett objekt (Y)
Y	Symbol som anger objekt, organism eller materie som är skild från den agerande organismen
Z	Extern observatör till en subjekt-objekt-relation

Statistiska symboler

A	Precisionsfaktor i ANOVA som anger variationsorsak aspekt (perception och värdering)
a_1	Nivå under A som anger perception
a_2	Nivå under A som anger värdering
α	Sannolikhet för att förkasta H_0 när H_0 är sann (sannolikhet för fel av typ I)
α_C	Cronbachs alfa, som är ett mått på reliabiliteten i ett enskilt variabelområde. Koefficienten anger reliabilitetens nedre gräns
b	Beteckning för en kanonisk variabel som hänför sig till bedömare
χ^2	(grekisk chi-bokstav). Slumpmässig variabel som är chi-kvadratfördelad
df	"Degree(s) of freedom", frihetsgrad(er)
e	Antal nivåer under en faktor (A) minus 1
ES	Effektstorlek som anger ett specifikt populationsvärde vilket avviker från noll, dvs $ES = 0$ när H_0 accepteras och $ES \neq 0$ när H_0 förkastas
F	F-kvot (F-test) som är beräknad på stickprovsvärden. Den anger den slumpmässigt fördelade variabeln "F"
F_{df_1, df_2}	Slumpmässig variabel som är fördelad enligt F med df_1 och df_2 frihetsgrader
f	"Effect size index". Den anger standardavvikelser för standardiserade medelvärden
g	Beteckning för styrkan i de erhållna testresultaten
H	Experimentell faktor i ANOVA som anger variationsorsak traditionell handledning (dyadisk konfrontation)
H_0	Hypotes som testas
H_1	Alternativ hypotes – hypotesen som accepteras när H_0 förkastas
h_1	Nivå under H som anger traditionell handledning
h_2	Nivå under H som anger ej traditionell handledning

H_i^2	Summan av de kvadrerade faktorladdningarna för variabel "i" i en population
h_i^2	Summan av de kvadrerade faktorladdningarna för variabel "i" i ett stickprov
$\bar{h}^2$	Medelvärde för h_i^2 i ett enskilt variabelområde
I	Individfaktor i ANOVA
i	Allmänt löpande indexbeteckning som anger den i:te observationen bland n distinkta observationer
J	Allmän beteckning för en variationsorsak
k	Allmän beteckning som anger klasser
l	Beteckning för en kanonisk variabel, som hänför sig till lärarkandidater
m	Aritmetiskt medelvärde
MS	"Mean Squares", medelkvadratsummor i ANOVA
N	Antal individer i en population från vilken ett stickprov har dragits
n	Allmän beteckning för ett antal observationer
ω^2	(grekisk omega-bokstav). Hays' omega i kvadrat. Den anger den skattade relativa eller proportionella reduktionen i populationsvariansen för den beroende variabeln när den oberoende variabeln är given
$\hat{\omega}^2$	Stickprovsskattning av den proportionella reduktionen i variansen för den beroende variabeln när den oberoende variabeln är given – ett mått på sambandsstyrkan mellan designens olika faktorer och mätinstrumentet
P	Allmän beteckning för en korrelationskoefficient i en population
P_I	Intraklasskorrelationskoefficient för en population
PV	Proportionsvarians som förklaras av en enskild variationsorsak
R_c	Kanonisk korrelationskoefficient i en population
R_i^2	Multipel korrelation i kvadrat för den i:te variabeln i en population
R_{ii}	Reliabilitet för den i:te variabeln i en individpopulation
r_c	Kanonisk korrelationskoefficient i ett stickprov
r_i^2	Multipel korrelation i kvadrat för den i:te variabeln i ett stickprov
r_{ii}	Reliabilitet för den i:te variabeln i ett stickprov
r_{ki}	Allmän beteckning för intraklasskorrelationer
r_n	Konfidensintervallens nedre gräns vid skattningen av r_{21}
r_s	Homogenitetsskattning för en summavariabel med hjälp av Spearman–Browns "prophecy" formel
$r_ö$	Konfidensintervallens övre gräns vid skattningen av r_{21}
$\bar{r}^2$	Medelvärde för r_i^2
r_{21}	Intraklasskorrelationskoefficient för 2 bedömare och 1 lektionstillfälle
$\bar{r}_{21}$	Genomsnittlig intraklasskorrelationskoefficient som ett approximativt mått på reliabiliteten i ett enskilt variabelområde
s	Standardavvikelse, beräknad på ett stickprov
SS	"Sum of Squares", summan av kvadraten i ANOVA
T	Experimentell faktor i ANOVA som anger variationsorsak externt förmedlad självkonfrontation via ITV/VR-tekniken
t_1	Nivå under T som anger externt förmedlad självkonfrontation via ITV/VR-tekniken

t_2	Nivå under T som anger ej externt förmedlad självkonfrontation via ITV/VR-tekniken
U	Experimentell faktor i ANOVA som anger variationsorsak undervisningstillfälle (ml_1, ml_2)
u_1	Nivå under U som anger ml_1
u_2	Nivå under U som anger ml_2
V	Precisionsfaktor i ANOVA som anger observationsvariablerna i skattnings- och värderingsformuläret F III
$v_1, \ldots, v_x$	Nivåer under V som anger variabel 1, . . . , x
x	Allmän beteckning för antal variabler

Referenser

Abbott, W. *An experimental study of self-confrontation in counseling.* Los Angeles, Calif.: University of Southern California, 1965. (Icke tryckt doktorsavhandling, Microfilms, 65-9959.)

Acheson, K. *The effects of feedback from television recordings and three types of supervisory treatment on selected teacher behaviors.* Stanford, Calif.: Stanford University, 1964. (Icke tryckt doktorsavhandling, Microfilms, 64-13, 542.)

Alger, I. & Hogan, P. The use of videotape recordings in conjoint marital therapy in private practice. Paper presented at the annual meeting of the American Psychiatric Association, Atlantic City, May, 1966.

Allen, D. W. & Ryan, K. *Micro-teaching.* Reading, Mass.: Addison-Wesley, 1969.

Ampex manual. Information for Ampex 5103, 7000 and 7003. Elk Grove Village, Ill.: Ampex Corporation, Consumer Products Division, 1969.

Anderson, T. W. *An introduction to multivariate statistical analysis.* New York: Wiley, 1958.

Bartlett, M. S. The statistical significance of canonical correlations. *Biometrika,* 1941, *32,* 29–38.

Bellack, A. A. Methods for observing classroom behavior of teachers and students. Paper written for discussion at the conference on methods of determining criteria for the evaluation of comprehensive schools. (Mimeographed.) Berlin: Pädagogisches Zentrum, November, 1968.

Benchley, R. *Benchley – or else!* London: Dobson, 1948.

Berger, E. M. The relation between expressed acceptance of self and expressed acceptance of others. *J. abnorm. soc. Psychol.,* 1952, *47,* 778–782.

Biddle, B. J. & Adams, R. S. *An analysis of classroom activities.* Columbia, Mo.: Center for Research in Social Behavior, University of Missouri, 1967.

Bierschenk, B. Television as a technical aid in education and in educational and psychological research: A bibliography. *Didakometry,* Nr 24, 1969.

Bierschenk, B. Att mäta subjekt-objekt-relationer i externt förmedlade självkonfrontationsprocesser via intern television: Presentation av ett kategorisystem. *Testkonstruktion och testdata,* Nr 6, 1971. (a)

Bierschenk, B. Att strukturera olika datanivåer med hjälp av faktoranalystekniker. *Kompendieserien,* Nr 12, 1971. (b)

Bierschenk, B. Television som tekniskt hjälpmedel i utbildning och pedagogisk-psykologisk forskning: En bibliografi. *Pedagogisk dokumentation,* Nr 2, 1971. (c)

Bierschenk, B. Television as a technical aid in education and in educational and psychological research: A bibliography (continued). *Didakometry,* Nr 29, 1971. (d)

Bierschenk, B. Television as a technical aid in education and in educational and psychological research: A bibliographic account of German literature. *Didakometry,* Nr 31, 1971. (e)

Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Bakgrund, experiment, experimentupplevelse. *Pedagogisk-psykologiska problem,* Nr 154, 1971. (f)
Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Analyser av lärarkandidaters självbedömning. *Pedagogisk-psykologiska problem,* Nr 160, 1972. (a)
Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Analyser av pedagogiska experters bedömningar. *Pedagogisk-psykologiska problem,* Nr 164, 1972. (b)
Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Analyser av lärarkandidaters självbedömning och pedagogiska experters bedömningar. *Pedagogisk-psykologiska problem,* Nr 165, 1972. (c)
Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Resultat, implikationer och rekommendationer. *Pedagogisk-psykologiska problem,* Nr 170, 1972. (d)
Bierschenk, B. Självkonfrontation via intern television i lärarutbildningen: Test och testdata. *Testkonstruktion och testdata,* Nr 12, 1972. (e)
Bierschenk, B. Att mäta subjekt-objekt-relationer i externt förmedlade självkonfrontationsprocesser via intern television: Presentation av ett kategorisystem. (Rev uppl.) *Testkonstruktion och testdata,* Nr 6, 1972. (f)
Bills, R., Macagnoni, V. & Elliot, R. *Student teacher personality change as a function of the personalities of supervising and cooperating teachers.* College of Education, University of Alabama. Washington D.C.: United States Office of Education, 1964.
Birdwhistell, R. L. *Introduction to kinesics.* Louisville, Ky.: Univer. of Louisville Press, 1953.
Bjerstedt, Å. Några interaktionsrelaterade variabler av potentiellt intresse för grupp-orienterad vuxenutbildning. *Pedagogisk-psykologiska problem,* Nr 12, 1965.
Bjerstedt, Å. Tele-auskultationer: Intern television och videobandning i lärarutbildningen. *Skola och Samhälle,* 1966, *47,* 179–201.
Bjerstedt, Å. CCTV and videorecordings as "observation amplifiers" in teacher training. *Education Television International,* 1967, *1,* 300–312. (a)
Bjerstedt, Å. Das ITV-Projekt: Internes Fernsehen in der Lehrerausbildung. *Didakometrie und Soziometrie,* Nr 2, 1967. (b)
Bjerstedt, Å. Interaction-oriented approaches to the assessment of student teachers. *J. Teach. Educ.,* 1967, *18,* 339–357. (c)
Bjerstedt, Å. *Rosenzweig Picture-Frustration Study: Vuxenversionen.* Svensk bearbetning. Stockholm: Skand. Testförlaget, 1968. (a)
Bjerstedt, Å. Videobandade simulatortest för analys av interaktionstendenser. *Pedagogisk Tidskrift,* 1968, *104,* 121–136. (b)
Bjerstedt, Å. (Red.) Aktuella forskningsprojekt vid institutionen. *Pedagogisk-psykologiska problem,* Nr 125, 1970. (a)
Bjerstedt, Å. *Den pedagogiska processen.* Lund: Gleerups, 1970. (b)
Bjerstedt, Å. & Sundgren, P. Interaction tendencies, personality, and teacher effectiveness. *Scand. J. educ. Res.,* 1968, *12,* 51–90.
Box, G. E. P. Non-normality and tests on variance. *Biometrika,* 1953, *40,* 318–335.
Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems. *Ann. Math. Statist.,* 1954, *25,* 290–302, 484–498.
Bracht, G. H. & Glass, G. V. The external validity of experiments. *Amer. educ. Res. J.,* 1968, *5,* 437–474.
Brusling, Ch. Microteaching (förkortat MU-projektet). *Information om skolforskning,* 1971:4.
Buckley, W. (Ed.) *Modern systems research for the behavioral scientist.* Chicago, Ill.: Aldine, 1968.

Burt, C. L. The stability of factors. *Brit. J. stat. Psychol.,* 1964, *17,* 177–180.
Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental design for research on teaching. I: Gage, N. L. (Ed.), *Handbook of research on teaching.* Chicago, Ill.: Rand McNally, 1963. Ss. 171–246.
Carlsson, G. *Social mobility and class structure.* Lund: Gleerups, 1958.
Cattell, R. B. & Warburton, F. W. *Objective personality and motivation tests. A theoretical introduction and practical compendium.* Chicago, Ill.: University of Illinois Press, 1967.
Cohen, J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1969.
Cohen, L. Functional dependence, exchanges, and power of influence. *Int. J. Educ. Sci.,* 1969, *3* (1), 47–51.
Combs, W. A. *The professional education of teachers.* Boston, Mass.: Allen & Bacon, 1965.
Combs, W. A. & Snygg, D. *Individual behavior. A perceptual approach to behavior.* (Rev. uppl.) New York: Harper, 1959.
Cooley, Ch. H. The social self. On the meaning of "I". I: Gordon, Ch. & Gergen, K. J. (Eds.), *The self in social interaction. Vol. I: Classic and contemporary perspectives.* New York: Wiley, 1968. Ss. 87–91.
Cooley, W. W. & Lohnes, P. R. *Multivariate data analysis.* New York: Wiley, 1971.
Coombs, Ph. H. *Utbildningens världskris. Utbildningsexplosionen i u-land och i-land.* (Övers. Sundmark, R.) Stockholm: Bonniers, 1968.
Cronbach, L. J. *Essentials of psychological testing.* New York: Harper, 1960.
Davis, Flora. The way we speak "body language". *New York Times Magazine,* May 31, 1970, 8–42.
De Lemos, Thereza. *An exploratory study of the effects of a feed-back condition on counseling interviews with seventh graders.* Milwaukee, Wisc.: University of Wisconsin, 1963. (Icke tryckt doktorsavhandling, Microfilms, 63-5740.)
Dewey, J. The relation of theory to practice in education. National society for the scientific study of education. *Third Yearbook,* 1904. Reprinted by Ass. Stud. Teach., State College of Iowa, Bull. 17, 1962.
Diggory, J. C. *Self-evaluation. Concepts and studies.* New York: Wiley, 1966.
Dixon, W. J. (Ed.) BMD Biomedical computer programs. *Automatic Computation,* Nr 2 och 3, 1970.
Eachus, H. T. *Self-confrontation for complex skill training: Review and analysis.* Columbus, O.: Wright-Patterson AFB, 1965.
Eboch, S. *A process and systems structure for the field of audiovisual communications.* Los Angeles, Calif.: University of Southern California, 1962. (Icke tryckt doktorsavhandling, Microfilms, 62-6049.)
Ekman, G. Subjektiv bedömning. I: Westerlund, G. (Utg.), *Människan och arbetet.* Stockholm: Strömbergs Förlag, 1950. Ss. 441–477.
Fanslow, W. V. *The effects of three modes of feedback bases on 35 mm time-lapse.* Stanford, Calif.: Stanford University, 1967. (Icke tryckt doktorsavhandling, Microfilms, 68-6413.)
Fisher, R. A. *Statistical methods for research workers.* (10:e uppl.) London: Oliver & Boyd, 1946.
Flanders, N. A. "Integrating theory and practice in teacher education", *Theoretical bases for professional laboratory experiences in teacher education.* Forty-fourth Ass. Stud. Teach. Yearbook. Dubuque, Ia.: W. C. Brown, 1965.

Gage, N. L., Runkel, P. J. & Chatterjee, B. B. *Equilibrium theory and behavior change: An experiment in feedback from pupils to teachers.* Urbana, Ill.: University of Illinois, August, 1960. (Mimeographed.)

Gerbner, G. m.fl. (Eds.) *The analysis of communication content. Developments in scientific theories and computer techniques.* New York: Wiley, 1969.

Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1954.

Handbok för handledare vid praktisk lärarutbildning. Klass- och ämneslärarlinjerna. Stockholm: Skolöverstyrelsen, 1967.

Hays, W. L. *Statistics.* New York: Holt, 1970.

Heil, L. M., Powell, M. & Feifer, J. *Characteristics of the teacher behavior and competency related to the achievement of different kinds of children in several elementary grades.* New York: Office of Testing and Research, Brooklyn Coll., 1960. (Mimeographed.)

Heinich, R. *Instructional technology and instructional management: A proposal for a new theoretical structure.* Los Angeles, Calif.: University of Southern California, 1968. (Icke tryckt doktorsavhandling, Microfilms, 67-17, 688.)

Hofstätter, P. R. (Hrsg.) *Das Fischer Lexikon: Psychologie.* Frankfurt am Main: Fischer Bücherei, 1957.

Holzman, P. S. On hearing and seeing oneself. I: Geertsma, R. H. & Mackie, J. B. (Eds.), *Studies in self-cognition: Self-observation in the behavioral sciences.* Baltimore, Md.: The Williams & Wilkins Company, 1969. Ss. 198–209.

Horst, P. *Factor analysis of data matrices.* New York: Holt, 1965.

Hotelling, H. The most predictable criterion. *J. educ. Psychol.,* 1935, *26,* 139–142.

Husén, T. & Henricsson, S. *USTM: Uppmärksamhet.* Stockholm: Skand. Testförlaget, 1950.

Jersild, A. T. *When teachers face themselves.* New York: Bureau of Publication, Teacher College, Columbia University, 1955.

Kagan, N., Krathwohl, D. R. & Miller, R. Stimulated recall in therapy using videotape – A case study. *J. counsel. Psychol.,* 1963, *10* (3), 237–243.

Kendall, M. G. *A course in multivariate analysis.* London: Charles Griffin & Co., 1965.

Kerber, J. *Study of selected student teachers and modifications of their self-concept through use of portable video tape recordings.* Detroit, Mich.: Wayne State University, 1967. (Icke tryckt doktorsavhandling, Microfilms, 68-2094.)

Kerlinger, F. N. *Foundations of behavioral research.* New York: Holt, 1967.

Klasslärarutbildningen, studieplaner m.m. Stockholm: Skolöverstyrelsen, 1968.

Larsson, B. Matematiskt språkbruk i pedagogiska forskningsrapporter: Några exempel på missbruk. *Pedagogisk-psykologiska problem,* Nr 140, 1971.

Lee, J. A. *Test pattern. Instructional television at Scarborough College, University of Toronto.* Toronto: University of Toronto Press, 1971.

Lewin, K. Feedback problems of social diagnosis and action. I: Buckley, W. (Ed.) *Modern systems research for the behavioral scientist.* Chicago, Ill.: Aldine, 1968. Ss. 441–444.

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

Linnman, N. & Wennerberg, B. *Grundskolans naturkunskap, årskurs 4–6.* Stockholm: Bergvalls, 1962.

Lord, F. M. & Novick, M. R. *Statistical theories of mental test scores.* Reading, Mass.: Addison-Wesley, 1968.

Malmquist, E. *Läs- och skrivsvårigheter hos barn. Analys och behandlingsmetodik.* Lund: Gleerups, 1969.

McDonald, F. J. & Allen, D. W. *Training effects of feedback and modeling procedures on teaching performance.* Technical Report, Stanford Center for Research and Development in Teaching. Stanford University, School of Education, No. 3, 1967.

Medley, D. M. & Mitzel, H. E. Measuring classroom behavior by systematic observation. I: Gage, N. L. (Ed.), *Handbook of research on teaching.* Chicago, Ill.: Rand Mc Nally, 1963. Ss. 247–328.

Miller, G. A., Galanter, E. & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

Mortensen, E. *A critical study of procedures and instrumentation used in observing and recording classroom interaction behavior.* New York: Teachers College Press, Teachers College, Columbia University, 1971. (Icke tryckt doktorsavhandling.)

Müller, G. *Schul- und Studienfernsehen 1966.* Weinheim & Berlin: Beltz, 1966.

Nielsen, G. *Studies in self confrontation.* København: Munksgaard, 1962.

Olivero, J. *Video recordings as a substitute for live observations in teacher education.* Stanford, Calif.: Stanford University, 1964. (Icke tryckt doktorsavhandling, Microfilms, 65-2866.)

Orme, M. E. J. *The effects of modeling and feedback variables on the acquisition of a complex teaching strategy.* Stanford, Calif.: Stanford University, 1966. (Icke tryckt doktorsavhandling.)

Pepperman, C. *Effects upon counselors of the video-tape recorder used in the counseling practicum.* Philadelphia, Pa.: Pennsylvania State University, 1967. (Icke tryckt doktorsavhandling, Microfilms, 68-3556.)

Perlmutter, M. S. et al. Family diagnosis and therapy using videotape playback. *Am. J. Orthopsychiatry,* 1967, *37,* 900–905.

Roulx, K. *Physiological effects of tape recording on supervised counselors.* Amherst, Mass.: University of Massachusetts, 1967. (Icke tryckt doktorsavhandling, Microfilms, 67-6947.)

Rozeboom, W. W. *Foundations of the theory of prediction.* Homewood, Ill.: The Dorsey Press, 1966.

Rubenowitz, S. *Emotional flexibility-rigidity as a comprehensive dimension of mind. An empirical study of a construct, and its psychological and social implications.* P. A. Council Report, No. 34. Stockholm: Almqvist & Wiksell, 1963.

Salomon, G. & McDonald, F. J. *Pre- and post-test reactions to selfviewing one's teaching performance on videotape.* Stanford, Calif.: Stanford Center for Research and Development in Teaching, 1969. (Mimeographed.)

Schramm, W. Communication research in the United States. I: Schramm, W. (Ed.), *The science of human communication.* London: Basic Books, 1964.

Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

Stanley, J. *Improving experimental design and statistical analysis.* Chicago, Ill.: Rand McNally, 1967.

Steiner, L. D. & Rogers, E. D. Alternative responses to dissonance. *J. abnorm. soc. Psychol.:* 1963, *66,* 128–136.

Stickell, D. W. *A critical review of the methodology and results of research comparing televised and face-to-face instruction.* Philadelphia, Pa.: Pennsylvania State University, 1963. (Icke tryckt doktorsavhandling, Microfilms, 64-1419.)

Stoller, F. & Berger, M. M. Discussants: Video-tape papers. *Comparative Group Studies,* 1970, *1* (2), 177–190.

Stoller, F. H. Therapeutic concepts reconsidered in lights of videotape experience. *Comparative Group Studies,* 1970, *1* (2), 5–12.

Stroller, F. M. Group psychotherapy on television: An innovation with hospitalized patients. *Amer. Psychol.*, 1967, *22* (2), 158–163.
Sundgren, P. Rollpreferenser: Studier med ett attitydschema. *Pedagogisk-psykologiska problem*, Nr 26, 1966.
Sundgren, P. *Lärarroller och lärarlämplighet.* (Pedagogisk orientering och debatt, Nr 18.) Lund: UNISKOL, 1967.
Svenska AB Philips. *Ljusboken.* Stockholm. Philipshuset, Ljusavdelningen, 1966.
Swineford, E. J. A study of factors that affect teaching. *California J. educ. Res.*, 1963, *14*, 214–224.
Tatsuoka, M. M. *Multivariate analysis.* New York: Wiley, 1971.
Thomson, G. The maximum correlation of two weighted batteries (Hotelling's "Most predictable criterion"). *Brit. J. Psychol., Stat. Sec., Part 1*, 1947, *48*, 27–34.
Torgerson, W. S. *Theory and methods of scaling.* New York: Wiley, 1958.
Unruh, W. *The modality and validity of cues to lecture effectiveness.* Stanford, Calif.: Stanford University, 1967. (Icke tryckt doktorsavhandling, Microfilms, 68-6501.)
Walz, G. R. & Johnston, J. A. Counselors look at themselves on videotape. *J. counsel. Psychol.*, 1963, *10* (3), 232–236.
Ward, G. *Educating counselors by videotape using the interpersonal process recall technique.* East Lansing, Mich.: Michigan State University, 1965. (Icke tryckt doktorsavhandling, Microfilms, 66-6179.)
Warr, P. B. & Knapper, Ch. *The perception of people and events.* New York: Wiley, 1968.
Westrin, P. A. *WIT III.* Stockholm: Skand. Testförlaget, 1964.
Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.
Withall, J. & Lewis, W. W. Social interaction in the classroom. I: Gage, N. L. (Ed.) *Handbook of research on teaching.* Chicago, Ill.: Rand McNally, 1963. Ss. 683–714.
Wolf, W. *The expression of personality.* New York: Harper, 1943.
Woodworth, R. S. & Schlosberg, H. *Experimental psychology.* New York: Holt, 1954.
Wundt, W. *Grundriß der Psychologie.* (13 uppl.) Leipzig: Alfred Kröner Verlag, 1918.
Wylie, R. C. *The self concept. A critical survey of pertinent research literature.* Lincoln, Neb.: University of Nebraska Press, 1961.
Vägledning för inträdessökande till klasslärarlinje vid lärarhögskola. Stockholm: Skolöverstyrelsen, 15.4.1969.
Young, D. *The effectiveness of self-instruction in teacher education using modelling and video-tape feedback.* Stanford, Calif.: Stanford University, 1967. (Icke tryckt doktorsavhandling, Microfilms, 68-6518.)
Zajonc, R. The concepts of balance, congruity and dissonance. *Public Opinion Quarterly*, 1960, *24*, 280–296.

Published by the University of Lund, the Malmö School of Education, and the Society of Psychology and Pedagogy:

STUDIA PSYCHOLOGICA ET PAEDAGOGICA

SERIES ALTERA

1. Edlund, Sven: Diskussionen om begåvningsurvalet under reformations- och stormaktstiden. I. 1947.
2. Agrell, Jan: Nyinrättade professurer inom filosofiska fakulteten i Lund under 1800-talets första hälft. 1949.
3. Smith, Gudmund: Psychological Studies in Twin Differences. 1949.
4. Husén, Torsten: Om innebörden av psykologiska mätningar. Några bidrag till psykometrikens metodlära. 1949.
5. Smith, Gudmund: Interpretations of Behaviour Sequences. 1952.
6. Johannesson, Ingvar: Studier av sociala relationer mellan barn i folkskoleklasser. 1954.
7. Kragh, Ulf: The Actual-Genetic Model of Perception-Personality. 1955.
8. Bjerstedt, Åke: Interpretations of Sociometric Choice Status. 1956.
9. Werdelin, Ingvar: The Mathematical Ability. Experimental and Factorial Studies. 1958.
10. Werdelin, Ingvar: Geometrical Ability and the Space Factors in Boys and Girls. 1961.
11. Borg, Gunnar A. V.: Physical Performance and Perceived Exertion. 1962.
12. Lindell, Ebbe: Den svenska rättskrivningsmetodiken. 1964.
13. Bokander, Ingvar: Stereoscopically Induced Perceptual Conflict between Facial Photographs. A Discussion and Summary of Results. 1968.
14. Bjerstedt, Åke: Twelve Years of Educational and Psychological Research in Sweden: A Bibliography. 1968.
15. Larsson, Bernt: Bayes Strategies and Human Information Seeking. 1968.
16. Bjerstedt, Åke: Educational Technology. Instructional Programming and Didakometry. 1972.
17. Löfgren, Horst: The Measurement of Language Proficiency. 1972.
18. Bierschenk, Bernhard: Självkonfrontation via intern television i lärarutbildningen. 1972.

CWK GLEERUP
LUND/SWEDEN